

Méthodes numériques en finance

Nicolas Baradel

5 octobre 2023

Table des matières

Introduction	2
1 Mesure de l'erreur de Monte Carlo	4
2 Simulation de variables aléatoires	5
3 Méthodes de réduction de variance	7
3.1 Variable antithétique	7
3.2 Variable de contrôle	9
3.3 Fonction d'importance ou échantillonnage préférentiel	10
3.4 Méthode de stratification	11
4 Simulation de processus stochastiques	14
4.1 Simulation d'un mouvement brownien	14
4.2 Simulation d'une équation différentielle stochastique	15
4.2.1 Simulation exacte de processus d'Itô	15
4.2.2 Schéma d'Euler	17
4.2.3 Schéma de Milstein	18
4.3 Application à l'évaluation de produits dérivés	19
5 Méthodes des différences finies	19
5.1 Formule de Feynman-Kac	19
5.2 Schéma explicite	22
5.3 Schéma implicite	25
5.4 θ -schéma	25
6 Méthodes numériques en contrôle stochastique	26
6.1 Principe de programmation dynamique et caractérisation	26
6.2 Approche par Monte Carlo	27
6.3 Exemple - Gestion de portefeuille	27
6.3.1 Réduction de dimension	28
6.3.2 Exercice	29
6.4 Schéma explicite	30
6.5 Exemple - Gestion de portefeuille (suite)	31
6.5.1 Exercice	32
A Simulations - Méthode du rejet	33

B	Formule de Black Scholes	35
C	Exercice 11 - Prix des Zéros Coupon dans le modèle de Vašíček	35
D	Exercice 14 - Option asiatique - Calcul de l'espérance de X	35

Introduction

L'objectif est de calculer ou d'estimer des quantités de la forme :

$$\theta := \mathbb{E}[f(X)] = \int_{\mathbb{R}^d} f(x) d\mathbb{P}_X(x),$$

où

- X est une variable aléatoire à valeurs dans \mathbb{R}^d ,
- f est une fonction mesurable de \mathbb{R}^d dans \mathbb{R} ,
- \mathbb{P}_X est la loi de la variable aléatoire X ,
- $\mathbb{E}[|f(X)|] < +\infty$.

La méthode la plus simple est la **méthode de Monte Carlo**. Cette méthode consiste à simuler une suite (X_1, \dots, X_n) dont les composantes ne sont pas obligatoirement indépendantes, mais telles qu'une loi des grands nombres s'applique, c'est à dire que :

$$\hat{\theta}_n := \frac{1}{n} \sum_{k=1}^n f(X_k) \xrightarrow[n \rightarrow +\infty]{p.s.} \theta.$$

Dans le cas d'une suite i.i.d. (identiquement et indépendamment distribué), une telle convergence s'obtient grâce au théorème de la loi forte des grands nombres :

Théorème 0.1 (Loi forte des grands nombres). *Soit $(X_k)_{1 \leq k \leq n}$ une suite i.i.d. de variables aléatoires réelles telle que $\mathbb{E}[|X_1|] < +\infty$. Alors*

$$\frac{1}{n} \sum_{k=1}^n X_k \xrightarrow[n \rightarrow +\infty]{p.s.} \mathbb{E}(X_1).$$

On y associera un théorème central limite qui permet d'avoir une mesure de **l'erreur d'estimation**.

En finance, nous sommes souvent face à une variable aléatoire X qui est une semi-martingale d'Itô définie par une équation de la forme :

$$X_s^{t,x} = x + \int_t^s \mu(s, X_s^{t,x}) ds + \int_t^s \sigma(s, X_s^{t,x}) dW_s,$$

où :

- $x \in \mathbb{R}^d$,
- $\mu : [0, T] \times \mathbb{R}^d \rightarrow \mathbb{R}^d$,
- $\sigma : [0, T] \times \mathbb{R}^d \rightarrow S_d^+(\mathbb{R})$,
- W est un mouvement brownien,

et de telle sorte que l'équation ci-dessus ait une unique solution.

Ce cas ouvre la voie à une seconde méthode d'estimation de θ . La formule de Feynman Kac permet de relier le calcul d'intégral à une équation aux dérivées partielles qu'il est possible de résoudre numériquement.

Le cours se découpe de la façon suivante. Dans un premier temps nous mettrons en place les éléments qui permettent d'en déduire une mesure de l'erreur d'estimation par Monte Carlo (construction d'intervalles de confiance). Ensuite nous verrons des méthodes de simulation de variables aléatoires. Nous introduirons le concept de réduction de variance et nous verrons différents contextes. Nous simulerons ensuite des processus stochastiques afin d'appliquer les précédentes méthodes au calcul du prix d'actifs dérivés. Enfin, nous utiliserons les schéma numériques sur les équations aux dérivées partielles afin de calculer le prix d'actifs dérivés.

Les applications se feront avec Python.

1 Mesure de l'erreur de Monte Carlo

Il est essentiel d'être capable de mesurer l'erreur commise par une méthode d'estimation. Pour les méthodes de Monte Carlo, celle-ci repose souvent sur le théorème central limite.

Théorème 1.1 (Théorème central limite). *Soit $(X_k)_{1 \leq k \leq n}$ une suite i.i.d. de variables aléatoires réelles telle que $\mathbb{E}(X_1^2) < +\infty$. Alors*

$$\sqrt{n} \left(\frac{1}{n} \sum_{k=1}^n X_k - \mathbb{E}(X_1) \right) \xrightarrow[n \rightarrow +\infty]{\mathcal{L}} \mathcal{N}(0, \text{Var}(X_1)).$$

Grâce à ce théorème, nous pouvons définir un intervalle de confiance asymptotique.

Proposition 1.2. *Soit $\alpha \in]0, 1[$. Soit $\Phi^{-1}(\alpha)$ le quantile de niveau α de la loi normale. Soit :*

$$\begin{aligned} IC_{1-\alpha}^n[\mathbb{E}(X_1)] &= \left[\bar{X}_n \pm \Phi^{-1} \left(1 - \frac{\alpha}{2} \right) \frac{\sigma}{\sqrt{n}} \right] \\ &= \left[\bar{X}_n - \Phi^{-1} \left(1 - \frac{\alpha}{2} \right) \frac{\sigma}{\sqrt{n}}, \bar{X}_n + \Phi^{-1} \left(1 - \frac{\alpha}{2} \right) \frac{\sigma}{\sqrt{n}} \right], \end{aligned}$$

où $\bar{X}_n := \frac{1}{n} \sum_{k=1}^n X_k$ et $\sigma^2 := \text{Var}(X_1)$. Il s'agit d'un intervalle de confiance asymptotique sur $\mathbb{E}(X_1)$ de niveau $1 - \alpha$ c'est à dire :

$$\mathbb{P}(\mathbb{E}(X_1) \in IC_{1-\alpha}^n[\mathbb{E}(X_1)]) \xrightarrow[n \rightarrow +\infty]{} 1 - \alpha.$$

Le paramètre σ est généralement inconnu. On peut l'estimer avec :

$$\hat{\sigma} := \sqrt{\frac{1}{n-1} \sum_{k=1}^n (X_k - \bar{X}_n)^2}.$$

Remarque 1.3. *Dans Python, la fonction `std` du package `numpy` divise par n tandis que la fonction `sd` de R divise par $n - 1$.*

L'intervalle de confiance précédent reste valide asymptotiquement par application du théorème de Slutsky car $\hat{\sigma}$ est convergent.

Il est également possible de construire un intervalle de confiance sur $g(\bar{X}_n)$ avec g suffisamment régulière.

Théorème 1.4 (Méthode delta). *Soit $(X_k)_{1 \leq k \leq n}$ une suite de variables aléatoires réelles telle que*

$$\sqrt{n} (\bar{X}_n - \theta) \xrightarrow[n \rightarrow +\infty]{\mathcal{L}} \mathcal{N}(0, \sigma^2),$$

et $g : \mathbb{R} \rightarrow \mathbb{R}$ une fonction C^1 en θ et $g'(\theta) \neq 0$. Alors,

$$\sqrt{n} (g(\bar{X}_n) - g(\theta)) \xrightarrow[n \rightarrow +\infty]{\mathcal{L}} \mathcal{N}(0, \sigma^2 g'(\theta)^2).$$

Exemple 1.5. *Soit $(X_k)_{1 \leq k \leq n}$ une suite i.i.d. de loi de $\mathcal{P}(\lambda)$ avec $\lambda > 0$. On souhaite estimer $\mathbb{E}(X_1) = \lambda$.*

Alors $IC_{1-\alpha}^n[\lambda] = \left[\left(\sqrt{\bar{X}_n} \pm \frac{\Phi^{-1}(1-\frac{\alpha}{2})}{2\sqrt{n}} \right)^2 \right]$ est un intervalle de confiance asymptotique de

niveau $1 - \alpha$. On a ici stabilisé la variance : elle ne dépend plus de la quantité inconnue à estimer.

Démonstration. Par le théorème central limite :

$$\sqrt{n} (\bar{X}_n - \lambda) \xrightarrow[n \rightarrow +\infty]{\mathcal{L}} \mathcal{N}(0, \lambda).$$

On pose $g : x \mapsto \sqrt{x}$. On a $g'(x) = \frac{1}{2\sqrt{x}}$. En appliquant la méthode Delta, on en déduit :

$$\sqrt{n} \left(\sqrt{\bar{X}_n} - \sqrt{\lambda} \right) \xrightarrow[n \rightarrow +\infty]{\mathcal{L}} \mathcal{N} \left(0, \frac{1}{4} \right).$$

Il vient :

$$IC_{1-\alpha}^n[\sqrt{\lambda}] = \left[\sqrt{\bar{X}_n} \pm \frac{\Phi^{-1} \left(1 - \frac{\alpha}{2} \right)}{2\sqrt{n}} \right],$$

puis, comme g est strictement croissante sur \mathbb{R}_+ ,

$$IC_{1-\alpha}^n[\lambda] = \left[\left(\sqrt{\bar{X}_n} \pm \frac{\Phi^{-1} \left(1 - \frac{\alpha}{2} \right)}{2\sqrt{n}} \right)^2 \right] = \left[\left(\sqrt{\bar{X}_n} - \frac{\Phi^{-1} \left(1 - \frac{\alpha}{2} \right)}{2\sqrt{n}} \right)^2, \left(\sqrt{\bar{X}_n} + \frac{\Phi^{-1} \left(1 - \frac{\alpha}{2} \right)}{2\sqrt{n}} \right)^2 \right].$$

□

Exercice 1. Soit $(X_k)_{1 \leq k \leq n}$ une suite i.i.d. de loi $\mathcal{P}(\lambda)$ avec $\lambda > 0$. Comme dans l'exemple précédent, on souhaite estimer $\mathbb{E}(X_1) = \lambda$.

Montrez que $IC_{1-\alpha}^n[\lambda] = \left[\exp \left(\log(\bar{X}_n) \pm \frac{\Phi^{-1} \left(1 - \frac{\alpha}{2} \right)}{\sqrt{n\bar{X}_n}} \right) \right]$ est un intervalle de confiance asymptotique de niveau $1 - \alpha$.

2 Simulation de variables aléatoires

Il existe diverses méthodes de simulation de variables aléatoires réelles. L'une d'entre elles est universelle lorsqu'on connaît l'inverse de la fonction de répartition de la loi qu'on souhaite simuler. Définissons d'abord la fonction inverse.

Définition 2.1. Si on note F_X la fonction de répartition d'une variable aléatoire X , l'inverse généralisée de F_X , également appelée fonction quantile de la loi de X , est définie par

$$F_X^{\leftarrow}(\alpha) := \inf \{ x \in \mathbb{R} : F_X(x) \geq \alpha \}, \quad \alpha \in]0, 1[.$$

Lorsque F_X est inversible, c'est l'inverse habituel, c'est à dire que $F_X^{\leftarrow} = F_X^{-1}$.

On a le résultat fondamental suivant.

Théorème 2.2. Soit $U \sim \mathcal{U}([0, 1])$. Alors

$$F_X^{\leftarrow}(U) \stackrel{\mathcal{L}}{=} X.$$

Démonstration. Dans le cas où F_X est inversible, on a

$$F_{F_X^{-1}(U)}(x) = \mathbb{P}(F_X^{-1}(U) \leq x) = \mathbb{P}(U \leq F_X(x)) = F_X(x).$$

Les deux variables aléatoires ont la même fonction de répartition, elles ont donc la même loi. Pour le cas général, il faut montrer au préalable que

$$F_X^{\leftarrow}(\alpha) \leq x \iff \alpha \leq F_X(x),$$

puis la démonstration est identique. □

Exemple 2.3 (Loi de Bernoulli). Soit $X \sim \mathcal{B}(p)$ avec $p \in]0, 1[$. On a $F_X(x) = (1-p)\mathbf{1}_{[0,1[}(x) + \mathbf{1}_{[1,+\infty[}(x)$. On en déduit $F_X^{\leftarrow}(U) = \mathbf{1}_{[1-p,1]}(U) \stackrel{\mathcal{L}}{=} X$.

De l'exemple précédent, on déduit encore plus simplement que $\mathbf{1}_{[0,p]}(U) \stackrel{\mathcal{L}}{=} X$ car $U \stackrel{\mathcal{L}}{=} 1 - U$. De manière générale, pour simuler une loi discrète, il suffit de découper U dans des intervalles dont la taille est la probabilité de chaque évènement, et d'y associer l'évènement correspondant (ici 0 avec probabilité $1 - p$ et 1 avec probabilité p).

Exemple 2.4 (Loi exponentielle). Soit $X \sim \mathcal{E}(\lambda)$ avec $\lambda > 0$. On a $F_X(x) = (1 - e^{-\lambda x})\mathbf{1}_{\{x \geq 0\}}$. On en déduit $F_X^{\leftarrow}(U) = -\frac{1}{\lambda} \log(1 - U) \stackrel{\mathcal{L}}{=} X$.

Dans le cas continue, il suffit de calculer l'inverse classique et de l'appliquer. On peut également utiliser le fait que $U \stackrel{\mathcal{L}}{=} 1 - U$ et en déduire plus simplement que $-\frac{1}{\lambda} \log(U) \stackrel{\mathcal{L}}{=} X$.

Remarque 2.5. La fonction de répartition n'est pas toujours explicite. C'est le cas de la loi normale : celle-ci est approximée dans tous les langages de calcul scientifique (*norm.cdf* pour Python via le package *scipy.stats* et *pnorm* dans R, et respectivement *norm.ppf* et *qnorm* pour les inverses).

Exercice 2. Estimer via Monte Carlo les quantités suivantes :

- $\mathbb{E}[\sin(U)]$ où $U \sim \mathcal{U}([0, \pi])$
- $\mathbb{E}[\log(X)]$ où $X \sim \mathcal{E}(1)$
- $\mathbb{E}[(e^{-\frac{\sigma^2}{2} + \sigma Z} - 1)^+]$ où $Z \sim \mathcal{N}(0, 1)$ et $\sigma = 0.2$

On n'utilisera que des simulations de loi uniforme pour cet exercice. On utilisera $n = 10^6$ et on donnera un intervalle de confiance pour chacune des estimations.

Pour simuler une loi $\mathcal{U}([0, 1])$, dans Python on utilisera la fonction `random.random(n)` du package `numpy`, dans R on utilisera la fonction `runif(n)` où n est le nombre de simulations.

On pourra afficher les histogrammes de U , X et Z .

Exercice 3. Soit $(X, Y) \sim \mathcal{U}([-1, 1]^2)$.

- Montrez que $\mathbb{P}(X^2 + Y^2 \leq 1) = \frac{\pi}{4}$.
- En déduire un algorithme d'estimation de π par méthode de Monte Carlo et le mettre en place en Python.

Pour simuler une loi $\mathcal{U}([a, b])$, dans Python on utilisera la fonction `random.uniform(a, b, n)` du package `numpy`.

Proposition 2.6 (Méthode de Box-Muller). Soit $(U, V) \sim \mathcal{U}([0, 1]^2)$. Soient

$$X := \sqrt{-2 \log(U)} \cos(2\pi V), \quad Y := \sqrt{-2 \log(U)} \sin(2\pi V)$$

Alors X et Y sont deux variables aléatoires indépendantes de loi $\mathcal{N}(0, 1)$.

Démonstration. On utilise la caractérisation de la loi suivante : deux variables aléatoires A et B à valeurs dans \mathbb{R}^d ont la même loi si et seulement si, pour toute fonction continue bornée $f : \mathbb{R}^d \mapsto \mathbb{R}$,

$$\mathbb{E}[f(A)] = \mathbb{E}[f(B)].$$

Soit $(W, Z) \sim \mathcal{N}(0, 1)^{\otimes 2}$. Alors, avec le changement de variable $(w, z) = (r \cos(\theta), r \sin(\theta))$,

$$\begin{aligned} \mathbb{E}[f(W, Z)] &= \int_{\mathbb{R}} \int_{\mathbb{R}} f(w, z) \frac{1}{2\pi} e^{-\frac{w^2+z^2}{2}} dw dz \\ &= \int_0^{2\pi} \frac{1}{2\pi} \int_0^{+\infty} f(r \cos(\theta), r \sin(\theta)) r e^{-\frac{r^2}{2}} dr d\theta \\ &= \int_0^{2\pi} \frac{1}{2\pi} \int_0^{+\infty} f(\sqrt{s} \cos(\theta), \sqrt{s} \sin(\theta)) \frac{1}{2} e^{-\frac{s}{2}} ds d\theta. \end{aligned}$$

Par ailleurs, par l'Exemple 2.4, $S := -2 \log(U) \sim \mathcal{E}(\frac{1}{2})$ et $\Theta := 2\pi V \sim \mathcal{U}([0, 2\pi])$,

$$\begin{aligned} \mathbb{E}[f(X, Y)] &= \mathbb{E}\left[f(\sqrt{S} \cos(\Theta), \sqrt{S} \sin(\Theta))\right] \\ &= \int_0^{2\pi} \frac{1}{2\pi} \int_0^{+\infty} f(\sqrt{s} \cos(\theta), \sqrt{s} \sin(\theta)) \frac{1}{2} e^{-\frac{s}{2}} ds d\theta. \end{aligned}$$

□

Exercice 4. Implémentez la méthode de Box-Muller dans Python. On testera la normalité avec un test de Kolmogorov-Smirnov grâce à la fonction `kstest` du package `scipy.stats` (fonction `ks.test` dans R).

3 Méthodes de réduction de variance

Une fois la méthode de simulation choisie, on estime

$$\theta := \mathbb{E}[f(X)] = \int_{\mathbb{R}^d} f(x) d\mathbb{P}_X(x).$$

Dans le cas standard, on simule un échantillon $(X_1, \dots, X_n) \stackrel{i.i.d.}{\sim} \mathbb{P}_X$ puis on estime θ par :

$$\hat{\theta}_n := \frac{1}{n} \sum_{k=1}^n f(X_k).$$

L'estimateur est sans biais et sa précision est directement liée à sa variance qui est

$$\text{Var}(\hat{\theta}_n) = \frac{\sigma_{f(X)}^2}{n}.$$

L'objectif des méthodes de réduction de variance est de réduire $\sigma_{f(X)}^2$. C'est-à-dire définir une nouvelle suite de variables aléatoires i.i.d. $(Y_n)_{n \geq 1}$ telles que $\mathbb{E}(Y_1) = \theta$ et $\text{Var}(Y_1) < \sigma_{f(X)}^2$ sur laquelle prendre la moyenne.

3.1 Variable antithétique

Soit X une variable aléatoire telle que $X \stackrel{\mathcal{L}}{=} a - X$ (symétrie en $\frac{a}{2} \in \mathbb{R}$). En conséquence, $f(X) \stackrel{\mathcal{L}}{=} f(a - X)$. On pose l'estimateur antithétique :

$$\hat{\theta}_n^A := \frac{1}{n} \sum_{k=1}^n \frac{f(X_k) + f(a - X_k)}{2}.$$

Cet estimateur demeure sans biais et on a

$$\text{Var}(\hat{\theta}_n^A) = \frac{\sigma_{f(X)}^2 + \text{Cov}(f(X), f(a - X))}{2n}.$$

Lemme 3.1. *L'estimateur antithétique $\hat{\theta}_n^A$ est meilleur que l'estimateur $\hat{\theta}_n$.*

Démonstration. L'inégalité de Cauchy-Schwarz implique que $\text{Cov}(f(X), f(a - X)) \leq \sigma_{f(X)}^2$. Ainsi,

$$\text{Var}(\hat{\theta}_n^A) \leq \frac{\sigma_{f(X)}^2}{n} = \text{Var}(\hat{\theta}_n).$$

□

Il est à noter que l'ajout de la version antithétique implique le calcul de f à nouveau. Il est intéressant de pouvoir quantifier l'amélioration, car si cette dernière n'est pas réelle, on est moins efficace.

Lemme 3.2. *Soit f une fonction monotone. Dans ce cas*

$$\text{Var}(\hat{\theta}_n^A) \leq \frac{\sigma_{f(X)}^2}{2n}.$$

Démonstration. Il faut montrer que $\text{Cov}(f(X), f(a - X)) \leq 0$, ce qui sera admis. □

Remarque 3.3. *Dans les hypothèses du lemme 3.2, l'estimateur antithétique est en pratique toujours plus efficace. En effet, on a*

$$\text{Var}(\hat{\theta}_n^A) \leq \text{Var}(\hat{\theta}_{2n}).$$

La variance de $\hat{\theta}_{2n}$ est au mieux égale à celle de $\hat{\theta}_n^A$. En doublant l'échantillon pour l'estimateur classique, on évalue autant de fois la fonction f avec les deux estimateurs. Mais il est moins coûteux de calculer $a - X$ que de simuler une nouvelle fois X . De plus, la variance de $\hat{\theta}_n^A$ sera généralement strictement inférieure.

Exemple 3.4. *Les deux exemples les plus utiles de variables aléatoires antithétiques sont :*

- Si $U \sim \mathcal{U}([0, 1])$, $U \stackrel{\mathcal{L}}{=} 1 - U$;
- Si $X \sim \mathcal{N}(0, 1)$, $X \stackrel{\mathcal{L}}{=} -X$.

Exercice 5. *On estime à nouveau via Monte Carlo les quantités de l'exercice 2,*

- $\mathbb{E}[\sin(U)]$ où $U \sim \mathcal{U}([0, \pi])$,
- $\mathbb{E}[\log(X)]$ où $X \sim \mathcal{E}(1)$,
- $\mathbb{E}[(e^{-\frac{\sigma^2}{2} + \sigma Z} - 1)^+]$ où $Z \sim \mathcal{N}(0, 1)$ et $\sigma = 0.2$,

mais en utilisant cette fois des variables aléatoires antithétiques. On utilisera $n = 10^6$ et on donnera une estimation de l'écart-type pour chacune des estimations, qu'on comparera à la version précédente sans utilisation de variable antithétique. Pour la simulation d'une loi $\mathcal{N}(0, 1)$, on utilisera la fonction `random.randn` du package `numpy` (`rnorm` dans `R`). Comparez les écarts-types et conclure. Pour une loi normale non centrée réduite, il y a la fonction `random.normal`.

Remarque 3.5. L'exercice précédent met en évidence que si f est paire en $\frac{a}{2}$, le point de symétrie, la méthode est inefficace. On retrouve la même variance que pour $\hat{\theta}_n$. En effet, si $f(x) = f(a - x)$ pour tout $x \in \mathbb{R}$, alors

$$\hat{\theta}_n^A = \frac{1}{n} \sum_{k=1}^n \frac{f(X_k) + f(a - X_k)}{2} = \frac{1}{n} \sum_{k=1}^n f(X_k) = \hat{\theta}_n.$$

De l'autre côté, on peut remarquer que la réduction de variance maximale est atteinte pour f impaire en la symétrie, c'est à dire $-f(x) = f(a - x)$ pour tout $x \in \mathbb{R}$ (à une constante près). En effet, dans ce cas

$$\hat{\theta}_n^A = \frac{1}{n} \sum_{k=1}^n \frac{f(X_k) + f(a - X_k)}{2} = \frac{1}{n} \sum_{k=1}^n 0 = 0.$$

On peut généraliser cela. Le cas où la méthode est inefficace est lorsque la covariance entre $f(X)$ et $f(a - X)$ est maximale, c'est-à-dire lorsqu'ils sont liés linéairement positivement et qu'il existe $\alpha > 0$, $\beta \in \mathbb{R}$ tels que :

$$f(x) = \alpha f(a - x) + \beta.$$

Le cas pair en la symétrie est le cas $\alpha = 1, \beta = 0$. De l'autre côté, le cas où la méthode est la plus efficace est le cas où elle éliminerait toute la variance, lorsque la covariance est minimale, c'est-à-dire qu'il existe $\alpha < 0$, $\beta \in \mathbb{R}$ tels que :

$$f(x) = \alpha f(a - x) + \beta.$$

Le cas impair en la symétrie est le cas $\alpha = -1, \beta = 0$. Plus on se sera proche de cette forme, plus la méthode sera efficace.

3.2 Variable de contrôle

Soit $f(X)$ une variable aléatoire dont on souhaite calculer l'espérance. On suppose qu'on connaît l'espérance de $g(X)$ et on pose $m := \mathbb{E}[g(X)]$. Soit $a \in \mathbb{R}$. On pose l'estimateur :

$$\hat{\theta}_n^{C,a} = \frac{1}{n} \sum_{k=1}^n f(X_k) - a(g(X_k) - m).$$

Cet estimateur demeure sans biais et on a

$$\text{Var} \left(\hat{\theta}_n^{C,a} \right) = \frac{\sigma_{f(X)}^2 + a^2 \sigma_{g(X)}^2 - 2a \text{Cov}(f(X), g(X))}{n}.$$

Lemme 3.6. La plus petite variance de $\hat{\theta}_n^{C,a}$ est atteinte pour

$$a^* := \frac{\text{Cov}(f(X), g(X))}{\sigma_{g(X)}^2},$$

et dans ce cas

$$\text{Var} \left(\hat{\theta}_n^{C,a^*} \right) = \frac{\sigma_{f(X)}^2 - \frac{\text{Cov}(f(X), g(X))^2}{\sigma_{g(X)}^2}}{n} = \frac{\sigma_{f(X)}^2 (1 - \rho(f(X), g(X))^2)}{n}$$

Démonstration. Nous avons un polynôme du second degré en a , son minimum est atteint lorsque sa dérivée s'anulle, c'est à dire en

$$2a^* \sigma_{g(X)}^2 - 2Cov(f(X), g(X)) = 0,$$

puis on en déduit le résultat. □

Remarque 3.7. Le cas $a = 0$ revient au cas sans variable de contrôle. On a une amélioration stricte si $a^* \neq 0$ c'est à dire si $Cov(f(X), g(X)) \neq 0$.

Exercice 6. Soit $X \sim \mathcal{N}(0, 1)$. On souhaite estimer $\mathbb{E}(|X|)$. Donnez une estimation et son écart-type avec 10^6 simulations de Monte Carlo en utilisant la variable de contrôle X^2 . Comparez au cas sans utiliser de variable de contrôle.

Exercice 7. On estime à nouveau via Monte Carlo les quantités de l'exercice 2,

- $\mathbb{E}[\sin(U)]$ où $U \sim \mathcal{U}([0, \pi])$,
- $\mathbb{E}[\log(X)]$ où $X \sim \mathcal{E}(1)$,
- $\mathbb{E}[(e^{-\frac{\sigma^2}{2} + \sigma Z} - 1)^+]$ où $Z \sim \mathcal{N}(0, 1)$ et $\sigma = 0.2$,

mais en introduisant des variables de contrôle simples dont vous connaissez l'espérance (un polynôme au plus d'ordre 2 pour les deux premiers, et on rappelle que $\mathbb{E}(e^{sZ}) = e^{s^2/2}$ pour le troisième).

Pour simuler la loi exponentielle, on pourra utiliser la fonction `random.exponential` du package `numpy` (`rexp` dans `R`).

Remarque 3.8. Comme le met en évidence la formule de la variance de la méthode avec variable de contrôle, la méthode est maximale lorsque la corrélation est parfaite, minimale et nulle lorsqu'il n'y a pas de corrélation. Pour obtenir une bonne variable de contrôle, il s'agit donc de trouver une variable aléatoire proche (à une transformation affine près) dont on sait calculer l'espérance.

3.3 Fonction d'importance ou échantillonnage préférentiel

Pour rappel, on souhaite calculer :

$$\theta := \mathbb{E}[f(X)] = \int_{\mathbb{R}^d} f(x) d\mathbb{P}_X(x).$$

Si \mathbb{P}_X admet la densité f_X par rapport à la mesure μ , et si h est une densité par rapport à μ , alors

$$\theta = \mathbb{E}[f(X)] = \int_{\mathbb{R}^d} f(x) f_X(x) d\mu(x) = \int_{\mathbb{R}^d} \left[\frac{f(x) f_X(x)}{h(x)} \right] h(x) d\mu(x) = \mathbb{E}^h \left[\frac{f(X) f_X(X)}{h(X)} \right].$$

Ici, \mathbb{E}^h est l'opérateur d'espérance où X a pour densité h .

On pose l'estimateur

$$\hat{\theta}_n^{I,h} = \frac{1}{n} \sum_{k=1}^n \frac{f(X_k) f_X(X_k)}{h(X_k)},$$

où $(X_k)_{1 \leq k \leq n}$ est une suite i.i.d. de densité h . L'estimateur demeure sans biais et

$$Var \left(\hat{\theta}_n^{I,h} \right) = \frac{Var^h \left(\frac{f(X) f_X(X)}{h(X)} \right)}{n}.$$

Le cas le plus utile est lors du calcul d'évènement rare ou d'espérance où la variable aléatoire n'est pas nulle uniquement lors d'un évènement rare. Dans ce cas, on change la loi de telle sorte à ce que la variable aléatoire ne soit pas souvent nulle sous la densité h .

Exercice 8. *On souhaite calculer*

$$\theta := \mathbb{E} \left((e^{-\frac{\sigma^2}{2} + \sigma Z} - K)^+ \right),$$

avec : $\sigma = 0.2, K \in \{1, 2, 3\}$.

— Implémentez en Python une méthode avec un échantillonnage préférentiel avec h la densité d'une $\mathcal{N}(\mu, 1)$ où

$$\mu := \frac{\log(K)}{\sigma} + \frac{\sigma}{2}.$$

— La valeur de μ est choisie afin que la partie positive soit nulle si et seulement si $Z \leq \mu$ (et $\mathbb{P}^h(Z \leq \mu) = \frac{1}{2}$). Déduisez-en une amélioration simple de votre estimateur en utilisant une symétrie et comparez.

Remarque 3.9. *La méthode utilisée dans le dernier point de l'exercice précédent peut être vue comme une utilisation combinée de variables antithétiques ou comme application particulière des méthodes de stratification de la section suivante.*

3.4 Méthode de stratification

Soit $(A_i)_{1 \leq i \leq \kappa}$ une partition de \mathbb{R}^d . Alors

$$\theta := \mathbb{E}(f(X)) = \sum_{i=1}^{\kappa} \mathbb{P}(X \in A_i) \mathbb{E}(f(X) | X \in A_i).$$

On suppose connu $p_{A_i} := \mathbb{P}(X \in A_i)$ pour $1 \leq i \leq \kappa$ et on va estimer les espérances conditionnelles à un évènement. Soit $(n_{A_i})_{1 \leq i \leq \kappa}$ et n tels que

$$n = \sum_{i=1}^{\kappa} n_{A_i}.$$

On va allouer exactement n_{A_i} simulations afin d'estimer $\mathbb{E}(f(X) | X \in A_i)$ pour $1 \leq i \leq \kappa$. Soit $(X_k^i)_{1 \leq k \leq n_{A_i}}$ des variables aléatoires indépendantes de loi $\mathbb{P}_{X|X \in A_i}$, pour $1 \leq i \leq \kappa$. On pose l'estimateur :

$$\hat{\theta}_n^S := \sum_{i=1}^{\kappa} \frac{p_{A_i}}{n_{A_i}} \sum_{k=1}^{n_{A_i}} f(X_k^i).$$

L'estimateur demeure sans biais et, en posant $\sigma_{f(X)|A_i}^2 := \text{Var}(f(X) | A_i)$, sa variance est

$$\text{Var}(\hat{\theta}_n^S) = \sum_{i=1}^{\kappa} \frac{p_{A_i}^2 \sigma_{f(X)|A_i}^2}{n_{A_i}}.$$

Lemme 3.10. *Soit n le nombre total de simulations, le choix des $(n_{A_i})_{1 \leq i \leq \kappa}$ qui minimise la variance est*

$$n_{A_i} = n \frac{p_{A_i} \sigma_{f(X)|A_i}}{\sum_{i=1}^{\kappa} p_{A_i} \sigma_{f(X)|A_i}}.$$

On a alors

$$\text{Var}(\hat{\theta}_n^S) = \frac{1}{n} \left(\sum_{i=1}^{\kappa} p_{A_i} \sigma_{f(X)|A_i} \right)^2.$$

Démonstration. On souhaite minimiser en $(n_{A_i})_{1 \leq i \leq \kappa}$ la quantité

$$\text{Var} \left(\hat{\theta}_n^S \right) = \sum_{i=1}^{\kappa} \frac{p_{A_i}^2 \sigma_{f(X)|A_i}^2}{n_{A_i}}.$$

sous la contrainte $n = n_{A_1} + \dots + n_{A_\kappa}$. On introduit le langragien $\lambda \in \mathbb{R}$, la condition du premier ordre donne, pour tout $1 \leq i \leq \kappa$,

$$-\frac{p_{A_i}^2 \sigma_{f(X)|A_i}^2}{n_{A_i}^2} + \lambda = 0 \iff n_{A_i} = \frac{p_{A_i} \sigma_{f(X)|A_i}}{\sqrt{\lambda}}.$$

Ainsi $n_{A_i} \propto p_{A_i} \sigma_{f(X)|A_i}$, on en déduit le résultat en réinjectant la contrainte $n = n_{A_1} + \dots + n_{A_\kappa}$. La variance devient

$$\text{Var} \left(\hat{\theta}_n^S \right) = \frac{1}{n} \sum_{i=1}^{\kappa} p_{A_i}^2 \sigma_{f(X)|A_i}^2 \frac{\sum_{j=1}^{\kappa} p_{A_j} \sigma_{f(X)|A_j}}{p_{A_i} \sigma_{f(X)|A_i}} = \frac{1}{n} \left(\sum_{i=1}^{\kappa} p_{A_i} \sigma_{f(X)|A_i} \right)^2.$$

□

Lemme 3.11. *L'estimateur par stratification optimale $\hat{\theta}_n^S$ est meilleur que l'estimateur $\hat{\theta}_n$*

Démonstration. On a

$$\mathbb{E} \left(f(X)^2 \right) = \mathbb{E} \left(\sum_{i=1}^{\kappa} \mathbf{1}_{\{X \in A_i\}} f(X)^2 \right) = \sum_{i=1}^{\kappa} p_{A_i} \mathbb{E} \left(f(X)^2 \mid X \in A_i \right).$$

On en déduit, en utilisant des inégalités de convexité

$$\begin{aligned} n \text{Var} \left(\hat{\theta}_n \right) &= \mathbb{E} \left(f(X)^2 \right) - \mathbb{E} \left(f(X) \right)^2 \\ &= \sum_{i=1}^{\kappa} p_{A_i} \mathbb{E} \left(f(X)^2 \mid X \in A_i \right) - \left(\sum_{i=1}^{\kappa} p_{A_i} \mathbb{E} \left(f(X) \mid X \in A_i \right) \right)^2 \\ &\geq \sum_{i=1}^{\kappa} p_{A_i} \mathbb{E} \left(f(X)^2 \mid X \in A_i \right) - \sum_{i=1}^{\kappa} p_{A_i} \mathbb{E} \left(f(X) \mid X \in A_i \right)^2 \\ &\geq \sum_{i=1}^{\kappa} p_{A_i} \sigma_{f(X)|A_i}^2 \geq \left(\sum_{i=1}^{\kappa} p_{A_i} \sigma_{f(X)|A_i} \right)^2 = n \text{Var} \left(\hat{\theta}_n^S \right). \end{aligned}$$

□

Malheureusement, en pratique, on ne peut pas connaître les poids optimaux car on ne connaît pas les $\sigma_{f(X)|A_i}^2$.

Sans aucune information sur les $\sigma_{f(X)|A_i}^2$, on peut faire comme s'ils étaient constants et prendre :

$$n_{A_i} = n p_{A_i}.$$

Dans ce cas l'estimateur reste plus efficace que l'estimateur classique de Monte Carlo.

Lemme 3.12. *Lorsque le nombre de simulations par strate est :*

$$n_{A_i} = np_{A_i},$$

On a :

$$\text{Var}(\hat{\theta}_n^S) = \frac{1}{n} \sum_{i=1}^{\kappa} p_{A_i} \sigma_{f(X)|A_i}^2.$$

En particulier :

$$\text{Var}(\hat{\theta}_n^S) \leq \text{Var}(\hat{\theta}_n).$$

Démonstration.

$$\text{Var}(\hat{\theta}_n^S) = \sum_{i=1}^{\kappa} \frac{p_{A_i}^2 \sigma_{f(X)|A_i}^2}{n_{A_i}} = \frac{1}{n} \sum_{i=1}^{\kappa} p_{A_i} \sigma_{f(X)|A_i}^2.$$

Dans la démonstration du lemme 3.11, un résultat intermédiaire était que

$$\text{Var}(\hat{\theta}_n) \geq \frac{1}{n} \sum_{i=1}^{\kappa} p_{A_i} \sigma_{f(X)|A_i}^2,$$

ce qui permet de conclure. □

Remarque 3.13. *Il est possible en pratique de faire quelques simulations dans chaque strate, afin d'estimer les $\sigma_{f(X)|A_i}$ puis d'utiliser ces estimations pour en déduire le nombre de simulations à effectuer en utilisant le Lemme 3.11.*

Une difficulté demeure : pouvoir simuler les lois conditionnelles $X | X \in A_i$.

Exemple 3.14. *Soit U une loi uniforme sur $[0, 1]$. Soit $\kappa \in \mathbb{N}^*$ et*

$$A_i := \left[\frac{i-1}{\kappa}, \frac{i}{\kappa} \right], \quad 1 \leq i \leq \kappa.$$

Alors $U | U \in A_i \sim \mathcal{U}(A_i)$. On pose

$$U_i = \frac{i-1}{\kappa} + \frac{U}{\kappa}.$$

On a bien par construction $U_i \sim \mathcal{U}(A_i)$. De plus $p_{A_i} = \frac{1}{\kappa}$.

Dans l'exemple précédent, si on simule X via l'inverse de la fonction de répartition, c'est à dire $X = F_X^{\leftarrow}(U)$, il suffit de stratifier U . Il est également possible de choisir des intervalles pour X plutôt que U .

Exemple 3.15. *Soit F_X une fonction de répartition et on pose $X = F_X^{\leftarrow}(U)$ où $U \sim \mathcal{U}([0, 1])$. Soit $(a_i)_{1 \leq i \leq \kappa}$ une suite de réels (éventuellements égaux à $-\infty$ ou $+\infty$) strictement croissante telle que $\mathbb{P}(X \in [a_0, a_\kappa]) = 1$.*

$$A_i := [F_X(a_{i-1}), F_X(a_i)[$$

Alors, $U | U \in A_i \sim \mathcal{U}(A_i)$. On pose

$$U_i = F_X(a_{i-1}) + (F_X(a_i) - F_X(a_{i-1}))U.$$

On a bien par construction $U_i \sim \mathcal{U}(A_i)$. De plus $p_{A_i} = F_X(a_i) - F_X(a_{i-1})$. On a également :

$$F_X^{-1}(U_i) \stackrel{\mathcal{L}}{=} X | X \in [a_{i-1}, a_i], \quad 1 \leq i \leq \kappa.$$

Enfin, un cas particulier intéressant est lorsqu'on connaît l'espérance conditionnelle sur certains A_i .

Lemme 3.16. *On suppose qu'on connaît $m_i := \mathbb{E}(f(X) | X \in A_i)$ pour $i < \alpha$ avec $\alpha \geq 2$.*

— *Si on ne tient pas compte des variances des classes, on pose*

$$n_{A_i} := \begin{cases} 0 & 1 \leq i < \alpha, \\ \frac{np_{A_i}}{\sum_{j=\alpha}^{\kappa} p_{A_j}} & \alpha \leq i \leq \kappa. \end{cases}$$

Puis on pose

$$\hat{\theta}_n^S := \sum_{i < \alpha} p_{A_i} m_i + \sum_{i=\alpha}^{\kappa} \frac{p_{A_i}}{n_{A_i}} \sum_{k=1}^{n_{A_i}} f(X_k^i).$$

Alors

$$\text{Var}(\hat{\theta}_n^S) = \sum_{i=\alpha}^{\kappa} \frac{p_{A_i}^2 \sigma_{f(X)|A_i}^2}{n_{A_i}} = \frac{\sum_{i=\alpha}^{\kappa} p_{A_i}}{n} \sum_{i=\alpha}^{\kappa} p_{A_i} \sigma_{f(X)|A_i}^2.$$

— *Si on tient compte des variances des classes, on pose*

$$n_{A_i} := \begin{cases} 0 & 1 \leq i < \alpha, \\ n \frac{p_{A_i} \sigma_{f(X)|A_i}}{\sum_{i=\alpha}^{\kappa} p_{A_i} \sigma_{f(X)|A_i}} & \alpha \leq i \leq \kappa. \end{cases}$$

Exercice 9. *On souhaite calculer*

$$\theta := \mathbb{E}\left(\left(e^{-\frac{\sigma^2}{2} + \sigma Z} - K\right)^+\right),$$

avec : $\sigma = 0.2, K \in \{1, 2, 3\}$.

Proposez un estimateur avec $\kappa = 2$ strates et $\kappa = 3$ strates (pour ce dernier cas, on posera $a_2 = a_1 + 1$) et implémentez sous Python.

4 Simulation de processus stochastiques

4.1 Simulation d'un mouvement brownien

L'objectif est de simuler un mouvement brownien $(W_t)_{0 \leq t \leq T}$. Toutefois, il ne sera pas possible numériquement d'avoir une trajectoire sur $[0, T]$. L'objectif est de simuler seulement pour des instants strictement croissants $(t_j)_{0 \leq j \leq m}$ dans $[0, T]$ avec $t_0 = 0$ et $t_m = T$:

$$(W_{t_j})_{0 \leq j \leq m}.$$

Rappelons la définition d'un mouvement brownien.

Définition 4.1 (Mouvement brownien). *Un processus W est un mouvement brownien si, en posant $\mathcal{F}_t := \sigma(W_s, s \leq t)$,*

- $W_0 = 0$ p.s.
- $t \mapsto W_t$ est continue p.s.
- $W_t - W_s \perp \mathcal{F}_s$ pour $t \geq s$,
- $W_t - W_s \sim \mathcal{N}(0, t - s)$ pour $t \geq s$.

Soit $(Z_j)_{0 \leq j \leq m}$ une suite de variables aléatoires indépendantes telle que

$$\begin{aligned} Z_0 &= 0, \\ Z_j &\sim \mathcal{N}(0, t_j - t_{j-1}), \quad 1 \leq j \leq m. \end{aligned}$$

On définit :

$$\begin{aligned} W_0^m &:= Z_0, \\ W_{t_{j+1}}^m &:= W_{t_j}^m + Z_{j+1}, \quad 0 \leq j \leq m-1, \end{aligned}$$

c'est-à-dire

$$W_{t_j}^m = \sum_{k=0}^j Z_k, \quad 0 \leq j \leq m.$$

Proposition 4.2. *Soit $(W_t)_{t \in [0, T]}$ un mouvement brownien. Pour toute suite $(t_j)_{0 \leq j \leq m}$,*

$$(W_{t_j})_{0 \leq j \leq m} \stackrel{\mathcal{L}}{=} (W_{t_j}^m)_{0 \leq j \leq m}.$$

En général, on prendra un pas de temps constant $\Delta t := t_j - t_{j-1}$ pour tout $1 \leq j \leq m$.

Remarque 4.3. *Le mouvement brownien est symétrique, c'est à dire qu'on a*

$$W \stackrel{\mathcal{L}}{=} -W.$$

Cette propriété pourra être utilisée comme réduction de variance par variable antithétique.

Exercice 10. *Soit $(W_t)_{t \geq 0}$ un mouvement brownien.*

- *Simulez $n = 10^4$ trajectoires avec un pas de temps de $\Delta t = 0,01$ sur $[0, T]$ (on choisira $T = 1$).*
- *En déduire un estimateur de $\mathbb{E} [\sup_{0 \leq t \leq T} W_t]$ et donnez sa précision (sans le biais).*
- *Utilisez $\Delta t = 10^{-4}$ et comparez. On utilisera ce pas pour la suite de l'exercice.*
- *Proposez un estimateur qui s'appuie sur un processus antithétique et comparez la précision.*
- *Proposez un estimateur qui s'appuie sur une variable de contrôle et comparez la précision : on utilisera W_T et $\int_0^T W_s ds$.*

Remarque 4.4. *L'exercice précédent met en évidence que l'usage d'une suite discrète de la trajectoire du mouvement brownien peut entraîner un biais dans le calcul de l'espérance, en particulier quand celle-ci dépend de toute la trajectoire. On remarque que les méthodes de réduction de variances apprises pour les variables aléatoires réelles, en particulier ici antithétique et de contrôle, s'appliquent sans difficulté au cas des processus.*

4.2 Simulation d'une équation différentielle stochastique

4.2.1 Simulation exacte de processus d'Itô

Soit le processus $(X_t)_{t \geq 0}$ solution de

$$X_t = x + \int_0^t \mu(s, X_s) ds + \int_0^t \sigma(s, X_s) dW_s,$$

où $\mu : [0, T] \times \mathbb{R} \rightarrow \mathbb{R}$ et $\sigma : [0, T] \times \mathbb{R} \rightarrow \mathbb{R}$ sont des fonctions Lipschitz en la seconde variable uniformément par rapport à la première, à croissance au plus linéaire en la seconde variable uniformément par rapport à la première, et W est un mouvement brownien. Soit $\mathcal{F}_t := \sigma(W_s, s \leq t)$. Alors le processus X est \mathcal{F} -adapté et Markovien relativement à \mathcal{F} , c'est à dire

$$\mathcal{L}((X_u)_{u \geq t} \mid \mathcal{F}_t) = \mathcal{L}((X_u)_{u \geq t} \mid X_t)$$

On pose :

$$\begin{aligned} X_0^m &:= x, \\ X_{t_j}^m &:= Z_j, \quad 1 \leq j \leq m, \end{aligned}$$

où Z_j suit la loi $\mathcal{L}(X_{t_j}^m \mid X_{t_{j-1}}^m)$.

Proposition 4.5. *Pour toute suite $(t_j)_{0 \leq j \leq m}$,*

$$(X_{t_j})_{0 \leq j \leq m} \stackrel{\mathcal{L}}{=} (X_{t_j}^m)_{0 \leq j \leq m}.$$

Exemple 4.6 (Processus de Vašíček). *Un processus de Vašíček est la solution de l'équation différentielle stochastique*

$$X_t = x + \int_0^t -a(X_s - \mu)ds + \int_0^t \sigma dW_s,$$

avec $a, \sigma > 0$ et $\mu \in \mathbb{R}$. Le processus se réécrit :

$$X_t = \mu + (x - \mu)e^{-at} + \sigma e^{-at} \int_0^t e^{as} dW_s.$$

En particulier, on a, en loi, si pour tout j , $\Delta t = t_{j+1} - t_j$,

$$X_{t_{j+1}} = \mu + (X_{t_j} - \mu)e^{-a\Delta t} + \sigma e^{-a\Delta t} \int_0^{\Delta t} e^{as} dW_s.$$

Ceci implique

$$X_{t_{j+1}} \mid X_{t_j} \sim \mathcal{N} \left(\mu + (X_{t_j} - \mu)e^{-a\Delta t}, \frac{\sigma^2}{2a} (1 - e^{-2a\Delta t}) \right).$$

Démonstration. □

Exercice 11 (Estimation d'un Zéro Coupon). *Soit $(r_t)_{t \geq 0}$ un processus de Vašíček qui représente le taux instantané sous la probabilité risque-neutre \mathbb{Q} (on se place en marché complet en absence d'opportunité d'arbitrage). Soit $P(t, T)$ la valeur d'un zéro coupon de maturité T à la date t . On rappelle que*

$$\left(P(t, T) e^{-\int_0^t r_s ds} \right)_{t \geq 0}$$

est une martingale sous \mathbb{Q} . En conséquence,

$$P(t, T) = \mathbb{E}^{\mathbb{Q}} \left(e^{-\int_t^T r_s ds} \mid \mathcal{F}_t \right).$$

On se place en date 0 et on observe $r_0 = 0.01$. On pose $a = 0.25, \mu = 0.025, \sigma = 0.0075, T = 2$. On souhaite estimer $P(0, T)$ et on s'appuiera sur une simulation exacte de la trajectoire. On utilisera $n = 10^4$ et $\Delta t = T \times 10^{-3}$. On calculera à chaque fois l'erreur sous forme d'un intervalle de confiance à 95% (sans le biais).

- Proposez une méthode de Monte Carlo simple pour estimer $P(0, t)$ sans méthode de réduction de variance.
- Proposez une méthode de réduction de variance par variable antithétique.
- Proposez une méthode de réduction de variance par variable de contrôle.

Remarque : dans le cas simple de cet exercice, il est possible de calculer explicitement $P(0, T)$, voir en Annexes.

Remarque 4.7. Les méthodes de réduction de variance précédentes offrent d'excellents résultats. En effet, on peut remarquer que l'intégrale du processus $(r_t)_{t \geq 0}$ est au voisinage de 0 et $\exp(x)$ est, dans ce cas, proche de $1 + x$.

4.2.2 Schéma d'Euler

Soit le processus $(X_t)_{t \geq 0}$ solution de

$$X_t = x + \int_0^t \mu(s, X_s) ds + \int_0^t \sigma(s, X_s) dW_s,$$

où $\mu : [0, T] \times \mathbb{R} \rightarrow \mathbb{R}$ et $\sigma : [0, T] \times \mathbb{R} \rightarrow \mathbb{R}$ sont des fonctions Lipschitz en la seconde variable uniformément par rapport à la première, à croissance au plus linéaire en la seconde variable uniformément par rapport à la première, et W est un mouvement brownien. On introduit le processus X^m aux temps $(t_j)_{0 \leq j \leq m}$ avec $X_0^m = x$ par :

$$X_{t_{j+1}}^m = X_{t_j}^m + \mu(t_j, X_{t_j}^m)(t_{j+1} - t_j) + \sigma(t_j, X_{t_j}^m)(W_{t_{j+1}} - W_{t_j}).$$

On a donc en loi :

$$X_{t_{j+1}}^m \mid \mathcal{F}_{t_j} \sim \mathcal{N} \left(X_{t_j}^m + \mu(t_j, X_{t_j}^m)(t_{j+1} - t_j), \sigma(t_j, X_{t_j}^m)^2(t_{j+1} - t_j) \right).$$

Le processus ci-dessus n'est défini qu'aux instants $(t_j)_{0 \leq j \leq m}$. Il est possible de le prolonger par interpolation linéaire sur $]t_j, t_{j+1}[$ par :

$$X_t^m := X_{t_j}^m + \frac{t - t_j}{t_{j+1} - t_j} (X_{t_{j+1}}^m - X_{t_j}^m).$$

Proposition 4.8. On pose $t_j = \frac{jT}{m}$, c'est-à-dire $\Delta t = \frac{T}{m}$. Pour tout $p \geq 1$, il existe une constante C_p telle que

$$\sup_{0 \leq t \leq T} \mathbb{E} [|X_t^m - X_t|^{2p}] + \mathbb{E} \left[\sup_{0 \leq j \leq m} |X_{t_j}^m - X_{t_j}|^{2p} \right] \leq \frac{C_p}{m^p}$$

De plus, pour tout $\beta < \frac{1}{2}$,

$$\lim_{m \rightarrow +\infty} m^\beta \sup_{0 \leq t \leq T} |X_t^m - X_t| = 0.$$

Démonstration. Admis. □

Exercice 12 (Estimation du prix d'un Call en volatilité locale). Soit $(S_t)_{t \geq 0}$ le processus de prix d'un actif sous la probabilité risque-neutre et caractérisé (admis) par l'équation différentielle stochastique :

$$dS_t = rS_t dt + \sigma S_t^\gamma dW_t,$$

avec $\frac{1}{2} \leq \gamma < 1$ et $S_0 > 0$. On admettra qu'on peut appliquer le schéma d'Euler. Le prix d'un Call de maturité $T > 0$ et de prix d'exercice $K > 0$ est

$$C(T) = e^{-rT} \mathbb{E} [(S_T - K)^+].$$

On pose $r = 0$, $\sigma = 0.2$, $\gamma = 0.8$, $S_0 = 1$, $K = 1.2$ et $T = 1$.

- Proposez une méthode de Monte Carlo simple pour estimer $C(T)$ sans méthode de réduction de variance.
- Proposez une méthode de réduction de variance par variable antithétique.
- Proposez une méthode de réduction de variance par variable de contrôle s'appuyant sur le modèle de Black-Scholes.

4.2.3 Schéma de Milstein

Soit le processus $(X_t)_{t \geq 0}$ solution de

$$X_t = x + \int_0^t \mu(s, X_s) ds + \int_0^t \sigma(s, X_s) dW_s,$$

où $\mu : [0, T] \times \mathbb{R} \rightarrow \mathbb{R}$ et $\sigma : [0, T] \times \mathbb{R} \rightarrow \mathbb{R}$ sont des fonctions Lipschitz en la seconde variable uniformément par rapport à la première, à croissance au plus linéaire en la seconde variable uniformément par rapport à la première, $x \in \mathbb{R}$ et W est un mouvement brownien. On introduit le processus X^m aux temps $(t_j)_{0 \leq j \leq m}$ avec $X_0^m = x$ par :

$$\begin{aligned} X_{t_{j+1}}^m &= X_{t_j}^m + \mu(t_j, X_{t_j}^m)(t_{j+1} - t_j) + \sigma(t_j, X_{t_j}^m)(W_{t_{j+1}} - W_{t_j}) \\ &\quad + \frac{1}{2} \sigma(t_j, X_{t_j}^m) \partial_x \sigma(t_j, X_{t_j}^m) ((W_{t_{j+1}} - W_{t_j})^2 - (t_{j+1} - t_j)). \end{aligned}$$

Le processus ci-dessus n'est défini qu'aux instants $(t_j)_{0 \leq j \leq m}$. Il est possible de le prolonger par interpolation linéaire sur $]t_j, t_{j+1}[$ par :

$$X_t^m := X_{t_j}^m + \frac{t - t_j}{t_{j+1} - t_j} (X_{t_{j+1}}^m - X_{t_j}^m).$$

Proposition 4.9. On pose $t_j = \frac{jT}{m}$. Pour tout $p \geq 1$, il existe une constante C_p telle que

$$\mathbb{E} \left[\sup_{0 \leq j \leq m} |X_{t_j}^m - X_{t_j}|^p \right] \leq \frac{C_p}{m^p}$$

De plus, pour tout $\beta < 1$,

$$\lim_{m \rightarrow +\infty} m^\beta \sup_{0 \leq t \leq T} |X_t^m - X_t| = 0.$$

Démonstration. Admis. □

Exercice 13 (Estimation du prix d'un Call en volatilité locale). Refaites l'exercice 12 en utilisant un schéma de Milstein et comparez.

4.3 Application à l'évaluation de produits dérivés

Exercice 14 (Option asiatique). Soit $(W_t)_{0 \leq t \leq T}$ un \mathbb{Q} - mouvement brownien sur $[0, T]$ et

$$S_t = S_0 e^{(r - \frac{\sigma^2}{2})t + \sigma W_t}.$$

On pose

$$C_T^a = \left(\frac{1}{T} \int_0^T S_t dt - K \right)^+.$$

On prendra $S_0 = 1, r = 0.01, \sigma = 0.2, T = 1$ et $K \in \{0.8, 1, 1.2\}$. On fixera $n = 10^4$ et $\Delta t = 10^{-4}$.

— Proposez une méthode de Monte Carlo afin d'estimer

$$C_0^a = e^{-rT} \mathbb{E}^{\mathbb{Q}} (C_T^a).$$

— On introduit la variable :

$$X := \left(\exp \left(\frac{1}{T} \int_0^T \log(S_t) dt \right) - K \right)^+.$$

Proposez une méthode d'estimation de C_0^a avec réduction de variance.

5 Méthodes des différences finies

5.1 Formule de Feynman-Kac

Soit le processus $(X^{t,x})_{\geq t}$ solution de

$$X^{t,x} = x + \int_t^\cdot \mu(s, X_s^{t,x}) ds + \int_t^\cdot \sigma(s, X_s^{t,x}) dW_s,$$

où $\mu : [0, T] \times \mathbb{R} \rightarrow \mathbb{R}$ et $\sigma : [0, T] \times \mathbb{R} \rightarrow \mathbb{R}$ sont des fonctions Lipschitz et W est un mouvement brownien.

Proposition 5.1. Soit $g : \mathbb{R} \rightarrow \mathbb{R}$ une fonction mesurable à croissance polynomiale. On pose :

$$v(t, x) = \mathbb{E} [g(X_T^{t,x})].$$

On a la relation

$$v(t, x) = \mathbb{E} [v(\tau, X_\tau^{t,x})], \quad 0 \leq t \leq \tau \leq T,$$

où τ est un temps d'arrêt. Si de plus v est de classe $C^{1,2}$, elle est l'unique solution de l'équation :

$$\begin{cases} \partial_t v(t, x) + \mu(t, x) \partial_x v(t, x) + \frac{1}{2} \sigma^2(t, x) \partial_{xx}^2 v(t, x) = 0 & (t, x) \in [0, T] \times \mathbb{R} \\ v(T, x) = g(x) & x \in \mathbb{R} \end{cases} \quad (1)$$

Démonstration. On pose

$$\tau := \inf\{s > t : |X_s^{t,x} - x| \geq 1\},$$

et $\tau_h := (t + h) \wedge \tau$. Comme v est régulière, par la formule d'Itô,

$$v(\tau_h, X_{\tau_h}^{t,x}) = v(t, x) + \int_t^{\tau_h} \partial_t v(s, X_s^{t,x}) ds + \int_t^{\tau_h} \mu(s, X_s^{t,x}) \partial_x v(s, X_s^{t,x}) ds \quad (2)$$

$$+ \int_t^{\tau_h} \sigma(s, X_s^{t,x}) \partial_x v(s, X_s^{t,x}) dW_s + \frac{1}{2} \int_t^{\tau_h} \sigma^2(s, X_s^{t,x}) \partial_{xx}^2 v(s, X_s^{t,x}) ds. \quad (3)$$

En appliquant l'espérance,

$$\mathbb{E} \left[\int_t^{\tau_h} \partial_t v(s, X_s^{t,x}) + \frac{1}{2} \sigma^2(s, X_s^{t,x}) \partial_{xx}^2 v(s, X_s^{t,x}) ds \right] + \mathbb{E} \left[\int_t^{\tau_h} \mu(s, X_s^{t,x}) \partial_x v(s, X_s^{t,x}) dW_s \right] = 0.$$

L'intégrande de la seconde intégrale est borné, l'intégrale stochastique est d'espérance nulle, il vient

$$\mathbb{E} \left[\frac{1}{h} \int_t^{\tau_h} \partial_t v(s, X_s^{t,x}) + \mu(s, X_s^{t,x}) \partial_x v(s, X_s^{t,x}) + \frac{1}{2} \sigma^2(s, X_s^{t,x}) \partial_{xx}^2 v(s, X_s^{t,x}) ds \right] = 0.$$

L'intégrande est borné, et comme $\frac{\tau_h - t}{h} \xrightarrow{p.s.} 1$, par le théorème de convergence dominée et le théorème de la valeur moyenne, on obtient le résultat. \square

Théorème 5.2 (Vérification). *Si φ est solution de l'équation aux dérivées partielles (1) alors $v = \varphi$.*

Démonstration. On introduit le temps d'arrêt

$$\tau_n := \inf\{s \geq t : |X_s^{t,x}| \geq n\}, \quad n \geq 1.$$

Par la formule d'Itô,

$$\begin{aligned} \varphi(\theta_n \wedge T) &= \varphi(t, x) + \int_t^{\theta_n \wedge T} \partial_t \varphi(s, X_s^{t,x}) + \frac{1}{2} \sigma^2(s, X_s^{t,x}) \partial_{xx}^2 \varphi(s, X_s^{t,x}) ds \\ &\quad + \int_t^{\theta_n \wedge T} \mu(s, X_s^{t,x}) \partial_x \varphi(s, X_s^{t,x}) dW_s. \end{aligned}$$

Comme φ vérifie l'équation aux dérivées partielles, la première intégrale est nulle. Grâce à la localisation, l'intégrande de la seconde intégrale est borné et d'espérance nulle, il vient

$$\mathbb{E} [\varphi(\theta_n \wedge T, X_{\theta_n \wedge T}^{t,x})] = \varphi(t, x).$$

Comme φ est à croissance polynomiale, pour $p \geq 2$,

$$\varphi(\theta_n \wedge T, X_{\theta_n \wedge T}^{t,x}) \leq C(1 + \sup_{t \leq s \leq T} |X_s^{t,x}|^p).$$

Sous les hypothèses du processus de diffusion, $\mathbb{E} [\sup_{t \leq s \leq T} |X_s^{t,x}|^p] < +\infty$ et enfin, comme $\theta_n \wedge T \xrightarrow{p.s.} T$, par le théorème de convergence dominée, on a le résultat par continuité puisque $\varphi(T, x) = g(x)$. \square

On a également le résultat plus général.

Proposition 5.3. Soient $f : [0, T] \times \mathbb{R} \rightarrow \mathbb{R}$ une fonction à croissance polynomiale en x uniformément en t et $r : [0, T] \times \mathbb{R} \rightarrow \mathbb{R}$ une fonction minorée. On définit :

$$v(t, x) = \mathbb{E} \left[\int_t^T e^{-\int_t^s r(u, X_u^{t,x}) du} f(s, X_s^{t,x}) ds + e^{-\int_t^T r(u, X_u^{t,x}) du} g(X_T^{t,x}) \right].$$

On a la relation

$$v(t, x) = \mathbb{E} \left[\int_t^\tau e^{-\int_t^s r(u, X_u^{t,x}) du} f(s, X_s^{t,x}) ds + e^{-\int_t^\tau r(u, X_u^{t,x}) du} v(\tau, X_\tau^{t,x}) \right], \quad 0 \leq t \leq \tau \leq T.$$

De plus, si v est de classe $C^{1,2}$, elle est l'unique solution de

$$\begin{cases} \partial_t v(t, x) + \mu(t, x) \partial_x v(t, x) + \frac{1}{2} \sigma(t, x)^2 \partial_{xx}^2 v(t, x) - r(t, x) v(t, x) + f(t, x) = 0 & (t, x) \in [0, T] \times \mathbb{R} \\ v(T, x) = g(x) & x \in \mathbb{R} \end{cases}$$

Démonstration. Admis. □

Théorème 5.4 (Principe de comparaison). On suppose que u est sous-solution et v est sur-solution de l'équation aux dérivées partielles sur $[0, T] \times \mathbb{R}$, c'est-à-dire, pour tout $(t, x) \in [0, T] \times \mathbb{R}$,

$$\partial_t u(t, x) + \mu(t, x) \partial_x u(t, x) + \frac{1}{2} \sigma(t, x)^2 \partial_{xx}^2 u(t, x) - r(t, x) u(t, x) + f(t, x) \leq 0 \quad (4)$$

$$\partial_t v(t, x) + \mu(t, x) \partial_x v(t, x) + \frac{1}{2} \sigma(t, x)^2 \partial_{xx}^2 v(t, x) - r(t, x) v(t, x) + f(t, x) \geq 0 \quad (5)$$

On suppose de plus que $u(T, \cdot) \geq v(T, \cdot)$, alors $u \geq v$.

Démonstration. Si r est minoré uniformément par une constante négative, on pourra poser $u'(t, x) = e^{\rho t} u(t, x)$ et $v'(t, x) = e^{\rho t} v(t, x)$ pour $(t, x) \in [0, T] \times \mathbb{R}$ et dans ce cas, u' et v' vérifient

$$\partial_t u'(t, x) + \mu(t, x) \partial_x u'(t, x) + \frac{1}{2} \sigma(t, x)^2 \partial_{xx}^2 u'(t, x) - (\rho + r(t, x)) u'(t, x) + e^{\rho t} f(t, x) \leq 0 \quad (6)$$

$$\partial_t v'(t, x) + \mu(t, x) \partial_x v'(t, x) + \frac{1}{2} \sigma(t, x)^2 \partial_{xx}^2 v'(t, x) - (\rho + r(t, x)) v'(t, x) + e^{\rho t} f(t, x) \geq 0 \quad (7)$$

Pour ρ assez grand, $\rho + r$ est minoré par une constante strictement positive, et le principe de comparaison sur u' et v' implique celui sur u et v . On pourra donc directement supposer que r est minoré par une constante strictement positive.

Par l'absurde, supposons qu'il existe $(t_0, x_0) \in [0, T] \times \mathbb{R}$ tel que

$$v(t_0, x_0) - u(t_0, x_0) > 0.$$

Comme u et v sont à croissance polynomiale, il existe $p \geq 1$ tel que

$$\limsup_{|x| \rightarrow +\infty} \frac{\sup_{t \leq T} (|u(t, x)| + |v(t, x)|)}{1 + |x|^p} = 0.$$

Pour $\varepsilon > 0$ et $\kappa \geq 0$, on introduit

$$\phi_{\varepsilon, \kappa}(t, x) = \varepsilon e^{\kappa t} (1 + x^{2p}).$$

Soit ε suffisamment petit tel que $v(t_0, x_0) - u(t_0, x_0) - \phi_{\varepsilon, \kappa}(t_0, x_0) > 0$. Alors il existe $(\hat{t}, \hat{x}) \in [0, T] \times \mathbb{R}$ tels que

$$\max_{(t, x) \in [0, T] \times \mathbb{R}} (v(t, x) - u(t, x) - \phi_{\varepsilon, \kappa}(t, x)) = v(\hat{t}, \hat{x}) - u(\hat{t}, \hat{x}) - \phi_{\varepsilon, \kappa}(\hat{t}, \hat{x}) > 0.$$

Comme $u(T, \cdot) \geq v(T, \cdot)$, $\hat{t} < T$. Les conditions d'optimalité du premier et du second ordre implique en particulier que

$$(\partial_t v - \partial_t u - \partial_t \phi_{\varepsilon, \kappa})(\hat{t}, \hat{x}) = 0, \quad (8)$$

$$(\partial_x v - \partial_x u - \partial_x \phi_{\varepsilon, \kappa})(\hat{t}, \hat{x}) = 0, \quad (9)$$

$$(\partial_{xx}^2 v - \partial_{xx}^2 u - \partial_{xx}^2 \phi_{\varepsilon, \kappa})(\hat{t}, \hat{x}) \leq 0. \quad (10)$$

En combinant la propriété de sous-solution de u avec celle de sur-solution de v , on obtient pour tout $(t, x) \in [0, T[\times \mathbb{R}$,

$$\partial_t(v - u)(t, x) + \mu(t, x)\partial_x(v - u)(t, x) + \frac{1}{2}\sigma(t, x)^2\partial_{xx}^2(v - u)(t, x) \geq r(t, x)(v - u)(t, x)$$

Au point (\hat{t}, \hat{x}) , en utilisant les conditions d'optimalités précédentes, il vient

$$\partial_t \phi_{\varepsilon, \kappa}(\hat{t}, \hat{x}) + \mu(\hat{t}, \hat{x})\partial_x \phi_{\varepsilon, \kappa}(\hat{t}, \hat{x}) + \frac{1}{2}\sigma(\hat{t}, \hat{x})^2\partial_{xx}^2 \phi_{\varepsilon, \kappa}(\hat{t}, \hat{x}) \geq r(\hat{t}, \hat{x})(v - u)(\hat{t}, \hat{x}).$$

En écrivant explicitement les dérivées de $\phi_{\varepsilon, \kappa}$, on remarque que pour κ assez grand, le membre de gauche devient négatif, ce qui est en contradiction avec le membre de droite qui est strictement positif. □

On remarque que le principe de comparaison implique également que, si une solution existe à l'équation aux dérivées partielles, elle est unique.

5.2 Schéma explicite

En réutilisant les fonctions introduites dans la section précédente, l'objectif est de résoudre une équation aux dérivées partielles de la forme

$$\begin{cases} \partial_t v(t, x) + \mu(t, x)\partial_x v(t, x) + \frac{1}{2}\sigma(t, x)^2\partial_{xx}^2 v(t, x) - r(t, x)v(t, x) + f(t, x) = 0 & (t, x) \in [0, T[\times \mathbb{R} \\ v(T, x) = g(x) & x \in \mathbb{R} \end{cases}$$

On supposera que les fonctions μ et σ vérifient les hypothèses habituelles, que r est minorée et que f admet au plus une croissance quadratique en x uniformément en t .

Nous avons besoin d'une maille discrète, sur laquelle nous allons approximer la fonction solution de l'équation aux dérivées partielles. On introduit Δt le pas de temps de discrétisation de $[0, T]$ tel que $T/\Delta t \in \mathbb{N}^*$. On introduit le maillage du temps :

$$\mathbf{T}^{\Delta t} := \{t_j := j\Delta t, j \leq T/\Delta t\}.$$

Pour $x \in \mathbb{R}$, nous sommes obligés de résoudre sur un espace borné. Soit Δx le pas d'espace et $c_-^{\Delta x} \leq c_+^{\Delta x}$ les bornes où nous allons résoudre numériquement l'équation aux dérivées partielles. On introduit le maillage de l'espace :

$$\mathbf{X}^{\Delta x} := \{x_j := j\Delta x, c_-^{\Delta x}/\Delta x \leq j \leq c_+^{\Delta x}/\Delta x\}.$$

On introduit les différences finies pour une fonction φ :

$$\partial_{\Delta t}\varphi(t, x) := \frac{\varphi(t + \Delta t, x) - \varphi(t, x)}{\Delta t}, \quad (11)$$

$$\partial_{\Delta x}^+\varphi(t, x) := \frac{\varphi(t, x + \Delta x) - \varphi(t, x)}{\Delta x}, \quad (12)$$

$$\partial_{\Delta x}^-\varphi(t, x) := \frac{\varphi(t, x) - \varphi(t, x - \Delta x)}{\Delta x}, \quad (13)$$

$$\partial_{\Delta x}^2\varphi(t, x) := \frac{\varphi(t, x + \Delta x) - 2\varphi(t, x) + \varphi(t, x - \Delta x)}{\Delta x^2}. \quad (14)$$

On évalue la condition terminale en posant $\varphi(T, x) = g(x)$, on pose des conditions en dehors du maillage en espace, au-delà des points $c_{-}^{\Delta x}$ et $c_{+}^{\Delta x}$ assez éloignés (on pourra y poser $\varphi(t, x) = g(x)$ ou une autre approximation) car les dérivées impliquent d'aller au-delà en espace. Le schéma explicite consiste à résoudre de manière rétrograde en t :

$$\partial_{\Delta t}\varphi(t, x) + \mu(t, x)\partial_{\Delta x}\varphi(t + \Delta t, x) + \frac{1}{2}\sigma(t, x)^2\partial_{\Delta x}^2\varphi(t + \Delta t, x) - r(t, x)\varphi(t + \Delta t, x) + f(t, x) = 0,$$

avec

$$\mu(t, x)\partial_{\Delta x}\varphi(t + \Delta t, x) := \begin{cases} \mu(t, x)\partial_{\Delta x}^+\varphi(t + \Delta t, x) & \text{si } \mu(t, x) \geq 0, \\ \mu(t, x)\partial_{\Delta x}^-\varphi(t + \Delta t, x) & \text{si } \mu(t, x) < 0. \end{cases}$$

La différence finie $\partial_{\Delta t}\varphi$ contient la seule valeur de φ évaluée en t , les autres sont en $t + \Delta t$ ce qui permet de d'obtenir facilement numériquement $\varphi(t, \cdot)$ connaissant $\varphi(t + \Delta, \cdot)$.

Définition 5.5 (Schéma explicite par différences finies).

- On pose $\varphi(T, x) = g(x)$ pour $x \in \mathbf{X}^{\Delta x}$,
- Sachant $\varphi(t + \Delta t, \cdot)$ pour un $t \in \mathbf{T}^{\Delta t}$, on déduit de l'équation des différences finies :

$$\begin{aligned} \varphi(t, x) &= \left(1 - \sigma^2(t, x)\frac{\Delta t}{\Delta x^2} - |\mu(t, x)|\frac{\Delta t}{\Delta x} - r(t, x)\Delta t \right) \varphi(t + \Delta t, x) \\ &+ \left(\frac{1}{2}\sigma^2(t, x)\frac{\Delta t}{\Delta x^2} + \mu^+(t, x)\frac{\Delta t}{\Delta x} \right) \varphi(t + \Delta t, x + \Delta x) \\ &+ \left(\frac{1}{2}\sigma^2(t, x)\frac{\Delta t}{\Delta x^2} + \mu^-(t, x)\frac{\Delta t}{\Delta x} \right) \varphi(t + \Delta t, x - \Delta x) \\ &+ f(t, x)\Delta t. \end{aligned}$$

Proposition 5.6. Si $(\Delta t, \Delta x) \rightarrow (0, 0)$ et tel qu'on ait toujours, pour $(\Delta t, \Delta x)$ suffisamment petit (condition de monotonie)

$$1 - \sigma^2(t, x)\frac{\Delta t}{\Delta x^2} - |\mu(t, x)|\frac{\Delta t}{\Delta x} - r(t, x)\Delta t \geq 0,$$

le schéma est convergent et φ converge simplement vers v .

Démonstration. Le schéma est monotone, car tous les coefficients devant les $\varphi(t + \Delta t, \cdot)$ sont positifs. Il est également stable, car la somme des coefficients devant les $\varphi(t + \Delta t, \cdot)$ est $1 - r(t, x)\Delta t$ et que r est minorée. Le schéma représente bien l'équation aux dérivées partielles qui admet une unicité forte. Avec cela, on peut démontrer la convergence du schéma. \square

Concrètement, si $(\Delta t, \Delta x) \rightarrow (0, 0)$ et que $\frac{\Delta t}{\Delta x^2} \rightarrow 0$, la condition de monotonie est satisfaite et le schéma est convergent. La difficulté est de s'assurer que $\frac{\Delta t}{\Delta x^2}$ soit suffisamment petit.

La monotonie peut être testée très facilement sur l'ensemble du schéma.

Remarque 5.7. Le schéma peut être vu comme une espérance, au facteur d'actualisation près. Si r est constant et $f = 0$, l'équation rétrograde pour la fonction de prix v peut s'écrire :

$$v(t, x) = e^{-r\Delta t} \mathbb{E} [v(t + \Delta t, X_{t+\Delta t}^{t,x})] = \int v(t + \Delta t, X_{t+\Delta t}^{t,x}(\omega)) d\mu(\omega),$$

où $\mu := e^{-r\Delta t} \mathbb{P}$. L'équation de la fonction φ du schéma explicite peut s'écrire

$$\mu_{\Delta x}^{t,x} := p(t, x)\delta_x + p^+(t, x)\delta_{x+\Delta x} + p^-(t, x)\delta_{x-\Delta x},$$

avec

$$\begin{aligned} p(t, x) &:= 1 - \sigma^2(t, x) \frac{\Delta t}{\Delta x^2} - |\mu(t, x)| \frac{\Delta t}{\Delta x} - r(t, x)\Delta t, \\ p^+(t, x) &:= \frac{1}{2}\sigma^2(t, x) \frac{\Delta t}{\Delta x^2} + \mu^+(t, x) \frac{\Delta t}{\Delta x}, \\ p^-(t, x) &:= \frac{1}{2}\sigma^2(t, x) \frac{\Delta t}{\Delta x^2} + \mu^-(t, x) \frac{\Delta t}{\Delta x}. \end{aligned}$$

Alors,

$$\varphi(t, x) = \int \varphi(t + \Delta t, X_{t+\Delta t}^{t,x}) d\mu_{\Delta x}^{t,x}(X_{t+\Delta t}^{t,x}),$$

et $p(t, x) + p^-(t, x) + p^+(t, x) = 1 - r(t, x)\Delta t$ qui est équivalent à $e^{-r\Delta t}$ lorsque que Δt tend vers 0.

Remarque 5.8. On peut écrire le schéma sous la forme matricielle :

$$\varphi(t, \cdot) = (I + A(t))\varphi(t + \Delta t, \cdot) + f(t, \cdot)\Delta t,$$

avec

$$A(t) := \begin{pmatrix} d(t, x_1) & d^+(t, x_1) & 0 & \dots & 0 & 0 & 0 \\ d^-(t, x_2) & d(t, x_2) & d^+(t, x_2) & \ddots & 0 & 0 & 0 \\ 0 & d^-(t, x_3) & d(t, x_3) & \ddots & 0 & 0 & 0 \\ \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \vdots \\ 0 & 0 & 0 & \ddots & d(t, x_{n-2}) & d^+(t, x_{n-2}) & 0 \\ 0 & 0 & 0 & \ddots & d^-(t, x_{n-1}) & d(t, x_{n-1}) & d^+(t, x_{n-1}) \\ 0 & 0 & 0 & \dots & 0 & d^-(t, x_n) & d(t, x_n) \end{pmatrix},$$

où

$$\begin{aligned} d(t, x) &:= p(t, x) - 1 = -\sigma^2(t, x) \frac{\Delta t}{\Delta x^2} - |\mu(t, x)| \frac{\Delta t}{\Delta x} - r(t, x)\Delta t, \\ d^+(t, x) &:= p^+(t, x), \\ d^-(t, x) &:= p^-(t, x). \end{aligned}$$

et où $\varphi(t, \cdot) = (\varphi(t, x_j))_j$ et $f(t, \cdot) = (f(t, x_j))_j$.

Exercice 15 (Estimation du prix d'un Call en volatilité locale). Refaites l'exercice 12 en utilisant le schéma explicite. On utilisera :

- $\Delta t = 0.00002$,
- $(c_-^{\Delta x}, c_+^{\Delta x}) = (0, 3)$,
- $\Delta x = 0.004$.

Testez la condition de monotonie. Observez le comportement du schéma avec $\Delta x = 0.002$.

5.3 Schéma implicite

Le schéma implicite consiste à évaluer les différences finies en espace à la date t au lieu de $t + \Delta t$, c'est à dire de résoudre l'équation :

$$\partial_{\Delta t}\varphi + \mu(t, x)\partial_{\Delta x}\varphi(t, x) + \frac{1}{2}\sigma(t, x)^2\partial_{\Delta x}^2\varphi(t, x) - r(t, x)\varphi(t, x) + f(t, x) = 0$$

On souhaite également résoudre le schéma de manière rétrograde, partant de $\varphi(T, x) = g(x)$. Toutefois, on ne peut pas en déduire immédiatement la valeur de $\varphi(t, x)$, mais on obtient un système linéaire des valeurs de $\varphi(t, \cdot)$ en fonction de celles de $\varphi(t + \Delta t, \cdot)$.

$$\begin{aligned} & \left(1 + \sigma^2(t, x)\frac{\Delta t}{\Delta x^2} + |\mu(t, x)|\frac{\Delta t}{\Delta x} + r(t, x)\Delta t \right) \varphi(t, x) \\ & + \left(-\frac{1}{2}\sigma^2(t, x)\frac{\Delta t}{\Delta x^2} - \mu^+(t, x)\frac{\Delta t}{\Delta x} \right) \varphi(t, x + \Delta x) \\ & + \left(-\frac{1}{2}\sigma^2(t, x)\frac{\Delta t}{\Delta x^2} - \mu^-(t, x)\frac{\Delta t}{\Delta x} \right) \varphi(t, x - \Delta x) \\ & = \varphi(t + \Delta t, x) + f(t, x)\Delta t \end{aligned}$$

L'équation ci-dessus conduit à la représentation matricielle :

$$(I - A(t))\varphi(t, \cdot) = \varphi(t + \Delta t, \cdot) + f(t, \cdot)\Delta t.$$

C'est-à-dire :

$$\varphi(t, \cdot) = (I - A(t))^{-1} (\varphi(t + \Delta t, \cdot) + f(t, \cdot)\Delta t),$$

avec $A(t)$ la matrice définie dans la Remarque 5.8.

Proposition 5.9. *Si $(\Delta t, \Delta x) \rightarrow (0, 0)$, le schéma est convergent et φ converge simplement vers v .*

5.4 θ -schéma

Le θ -schéma consiste à prendre une combinaison convexe du schéma explicite et implicite. C'est à dire de résoudre l'équation :

$$\begin{aligned} & \partial_{\Delta t}\varphi + \theta \left(\mu(t, x)\partial_{\Delta x}\varphi(t + \Delta t, x) + \frac{1}{2}\sigma(t, x)^2\partial_{\Delta x}^2\varphi(t + \Delta t, x) - r(t, x)\varphi(t + \Delta t, x) \right) \\ & + (1 - \theta) \left(\mu(t, x)\partial_{\Delta x}\varphi(t, x) + \frac{1}{2}\sigma(t, x)^2\partial_{\Delta x}^2\varphi(t, x) - r(t, x)\varphi(t, x) \right) + f(t, x) = 0 \end{aligned}$$

Matriciellement :

$$(I - (1 - \theta)A(t))\varphi(t, \cdot) = (I + \theta A(t))\varphi(t + \Delta t, \cdot) + f(t, \cdot)\Delta t,$$

c'est-à-dire

$$\varphi(t, \cdot) = (I - (1 - \theta)A(t))^{-1}(I + \theta A(t))\varphi(t + \Delta t, \cdot) + (I - (1 - \theta)A(t))^{-1}f(t, \cdot)\Delta t,$$

Le cas à privilégier est $\theta = \frac{1}{2}$ qui correspond au schéma de Crank-Nicholson.

6 Méthodes numériques en contrôle stochastique

Dans cette partie, on énonce les résultats *analogues* au cas sans contrôle qui permettent d'obtenir ensuite des schémas numériques de résolution.

6.1 Principe de programmation dynamique et caractérisation

Soit le processus $(X^{t,x,\phi})_{\geq t}$ solution de

$$X^{t,x,\phi} = x + \int_t^\cdot \mu(s, X_s^{t,x,\phi}, \phi_s) ds + \int_t^\cdot \sigma(s, X_s^{t,x,\phi}, \phi_s) dW_s,$$

où $\mu : [0, T] \times \mathbb{R} \times A \rightarrow \mathbb{R}$ et $\sigma : [0, T] \times \mathbb{R} \times A \rightarrow \mathbb{R}$ sont des fonctions Lipschitz et sous linéaires, W est un mouvement brownien, $A \subset \mathbb{R}$ et $\phi \in \mathcal{A}$ où \mathcal{A} est l'ensemble des processus \mathcal{F} -adapté de carré intégrable. Le processus ϕ est appelé le contrôle.

On rappelle le principe de programmation et l'équation aux dérivées partielles vérifiée dans le cas régulier.

Proposition 6.1. *Soit $g : \mathbb{R} \rightarrow \mathbb{R}$ une fonction mesurable à croissance polynomiale. On pose :*

$$v(t, x) = \sup_{\phi \in \mathcal{A}} \mathbb{E} \left[g(X_T^{t,x,\phi}) \right].$$

On a le principe de programmation dynamique

$$v(t, x) = \sup_{\phi \in \mathcal{A}} \mathbb{E} \left[v(\tau, X_\tau^{t,x,\phi}) \right], \quad 0 \leq t \leq \tau \leq T,$$

où τ est un temps d'arrêt. Si de plus v est de classe $C^{1,2}$, elle est l'unique solution de l'équation :

$$\begin{cases} \partial_t v(t, x) + \sup_{a \in A} \left[\mu(t, x, a) \partial_x v(t, x) + \frac{1}{2} \sigma^2(t, x, a) \partial_{xx}^2 v(t, x) \right] = 0 & (t, x) \in [0, T] \times \mathbb{R} \\ v(T, x) = g(x) & x \in \mathbb{R} \end{cases} \quad (15)$$

L'unicité et le contrôle optimal peuvent s'obtenir par un théorème de vérification.

Théorème 6.2. *On définit $\hat{a} : [0, T] \times \mathbb{R} \rightarrow A$ par*

$$\hat{a}[t, x] \in \arg \max_{a \in A} \left[\mu(t, x, a) \partial_x v(t, x) + \frac{1}{2} \sigma^2(t, x, a) \partial_{xx}^2 v(t, x) \right],$$

où on suppose que l'arg max n'est jamais vide. On pose

$$\hat{\phi}_s := \hat{a}[s, X_s^{t,x,\hat{\phi}}].$$

S'il existe $\gamma > 0$ tel que $\mathbb{E} \left[\sup_{s \in [t, T]} X_s^{t,x,\hat{\phi}} \right] < +\infty$ et que φ est une autre solution à croissance polynomiale d'au plus γ , alors $v = \varphi$ et la stratégie optimale est $\hat{\phi}$.

6.2 Approche par Monte Carlo

En réutilisant les fonctions introduites dans la section précédente, l'objectif est de s'appuyer sur le principe de programmation dynamique et de résoudre

$$\begin{cases} v(t, x) = \sup_{\phi \in \mathcal{A}} \mathbb{E} [v(\tau, X_\tau^{t,x,\phi})] & (t, x) \in [0, T[\times \mathbb{R} \\ v(T, x) = g(x) & x \in \mathbb{R} \end{cases}$$

Nous avons besoin d'une maille discrète, sur laquelle nous allons approximer la fonction solution. On introduit Δt le pas de temps de discrétisation de $[0, T]$ tel que $T/\Delta t \in \mathbb{N}^*$. On introduit le maillage du temps :

$$\mathbf{T}^{\Delta t} := \{t_j := j\Delta t, j \leq T/\Delta t\}.$$

Pour $x \in \mathbb{R}$, nous sommes obligés de résoudre sur un espace borné. Soit Δx le pas d'espace et $c_-^{\Delta x} \leq c_+^{\Delta x}$ les bornes où nous allons résoudre numériquement l'équation aux dérivées partielles. On introduit le maillage de l'espace :

$$\mathbf{X}^{\Delta x} := \{x_j := j\Delta x, c_-^{\Delta x}/\Delta x \leq j \leq c_+^{\Delta x}/\Delta x\}.$$

Remarque 6.3. Une difficulté de cette approche est qu'a priori, $X_s^{t,x,\phi}$ n'appartiendra pas à $\mathbf{X}^{\Delta x}$. On introduit $x \mapsto B(x)$ la fonction qui à $x \in \mathbb{R}$ associe les coins de l'hypercube autour de x (en dimension 1, il s'agit du point inférieur et du point supérieur). Pour tout $z \in B(x)$, on associe $z \mapsto p(z | x)$ les poids du barycentre formé par x dans $B(x)$ (en dimension 1, il s'agit des poids de l'interpolation linéaire).

Pour le contrôle, nous pouvons introduire un espace discret $\mathbf{A} \subset A$. Nous pouvons à présent définir le schéma.

Définition 6.4 (Schéma par Monte Carlo).

- On pose $\varphi(T, x) = g(x)$ pour $x \in \mathbf{X}^{\Delta x}$,
- Sachant $\varphi(t + \Delta t, \cdot)$ pour un $t \in \mathbf{T}^{\Delta t}$, on déduit de l'équation des différences finies :

$$\hat{a}[t, x] \in \arg \max_{a \in \mathbf{A}} \mathbb{E} \left[\sum_{z \in B(X_{t+\Delta t}^{t,x,\phi^a})} \varphi(t + \Delta t, z) p(z | X_{t+\Delta t}^{t,x,\phi^a}) \right]$$

et

$$\varphi(t, x) = \mathbb{E} \left[\sum_{z \in B(X_{t+\Delta t}^{t,x,\phi^{\hat{a}[t,x]}})} \varphi(t + \Delta t, z) p(z | X_{t+\Delta t}^{t,x,\phi^{\hat{a}[t,x]}}) \right].$$

6.3 Exemple - Gestion de portefeuille

Soit $(W_t^1)_{0 \leq t \leq T}$ un mouvement brownien sur $[0, T]$ et le prix de l'actif

$$dY_t = \mu Y_t dt + \sqrt{V_t} Y_t dW_t^1, \quad 0 \leq t \leq T.$$

Soit $(W_t^2)_{0 \leq t \leq T}$ un second mouvement brownien indépendant de W^1 . Le processus de variance qui apparait ci-dessus, V , est défini par

$$dV_t = -\alpha(V_t - \sigma^2)dt + \xi \sqrt{V_t} \left[\rho dW_t^1 + \sqrt{1 - \rho^2} dW_t^2 \right], \quad 0 \leq t \leq T.$$

Il s'agit du modèle de Henston. Pour plus de généralité, on introduit la notation $(Y_s^{t,y,v})_{t \leq s \leq T}$ qui est le processus Y sur $[t, T]$, partant de $Y_t = y > 0$ avec $V_t = v > 0$. De même on introduit $(V_s^{t,v})_{t \leq s \leq T}$ partant de $v > 0$ (ce processus est autonome et ne dépend pas de y).

On note $(X_s^{t,x,y,v})_{t \leq s \leq T}$ la valeur d'un portefeuille autofinçant partant de la richesse initiale $X_t^{t,x,y,v} = x \in \mathbb{R}$. On peut acheter et vendre des actions Y . On note ϕ_u la richesse investie en actions en date u . On considère les taux nuls, et par construction, la valeur du portefeuille ne varie que par l'investissement dans l'actif risqué. Le montant investi en actions est ϕ_u et, comme une action a un prix $Y_u^{t,y,v}$ en u , le nombre d'actions détenues en u est $\frac{\phi_u}{Y_u^{t,y,v}}$.

Par la condition d'autofinancement, la valeur du portefeuille contrôlé $(X_s^{t,x,y,v,\phi})_{t \leq s \leq T}$ est régie par

$$\begin{aligned} X_s^{t,x,y,v,\phi} &= x + \int_t^s \frac{\phi_u}{Y_u^{t,y,v}} dY_u^{t,y,v}, \\ &= x + \int_t^s \phi_u (\mu du + \sqrt{V_u^{t,v}} dW_u^1). \end{aligned} \quad (16)$$

Nous posons la fonction d'utilité U qui donne la valeur d'utilité terminale du portefeuille en fonction de la richesse :

$$U(x) = -\exp(-\eta x), \quad x \in \mathbb{R}.$$

Partant de (t, x, y, v) , on note \mathcal{A}^t l'ensemble des processus progressifs $(\phi_u)_{t \leq u \leq T}$ à valeurs dans \mathbb{R} tels que $X^{t,x,y,v,\phi}$ soit uniformément borné.

Soit $\phi \in \mathcal{A}^t$, on note

$$J(t, x, y, v, \phi) = \mathbb{E} \left[U(X_T^{t,x,y,v,\phi}) \right].$$

Par (16), on remarque que $X_T^{t,x,y,v,\phi}$ ne dépend pas de la condition initiale y . On peut définir J et X plus simplement :

$$J(t, x, v, \phi) = \mathbb{E} \left[U(X_T^{t,x,v,\phi}) \right].$$

L'objectif est de choisir $\phi \in \mathcal{A}^t$ afin de maximiser J . Cela permet de définir la fonction valeur

$$v(t, x, v) := \sup_{\phi \in \mathcal{A}^t} J(t, x, v, \phi) = \sup_{\phi \in \mathcal{A}^t} \mathbb{E} \left[U(X_T^{t,x,v,\phi}) \right].$$

Pour la suite, on supposera que v est de classe $C^{1,2}([0, T[, \mathbb{R} \times \mathbb{R}) \cap C([0, T], \mathbb{R} \times \mathbb{R})$.

6.3.1 Réduction de dimension

Certains problème peuvent être réduit en terme de dimension. Nous vevons déjà de le faire en remarquant que y n'a pas d'incidence. Dans les modèles à utilité exponentielle, il est fréquent que le contrôle optimal ne dépende pas de la richesse et de pouvoir réduire la dimension du problème.

Comme $X^{t,x,v,\phi} = x + X^{t,0,v,\phi}$ et que, par définition,

$$v(t, x, v) = \sup_{\phi \in \mathcal{A}^t} \mathbb{E} \left[-\exp(-\eta X_T^{t,x,v,\phi}) \right],$$

on obtient

$$v(t, x, v) = e^{-\eta x} \sup_{\phi \in \mathcal{A}^t} \mathbb{E} \left[-\exp(-\eta X_T^{t,0,v,\phi}) \right] = e^{-\eta x} v(t, 0, v), \quad (17)$$

On pose

$$v_o(t, v) := v(t, 0, v).$$

Pour θ un temps d'arrêt sur $[t, T]$, le principe de programmation dynamique s'écrit ici

$$v(t, x, v) = \sup_{\phi \in \mathcal{A}^t} \mathbb{E} \left[v(\theta, X_\theta^{t,x,v,\phi}, V_\theta^{t,v}) \right]. \quad (18)$$

De (17)-(18), on obtient facilement que

$$v_\circ(t, v) = \sup_{\phi \in \mathcal{A}^t} \mathbb{E} \left[e^{-\eta X_\theta^{t,0,v,\phi}} v_\circ(\theta, V_\theta^{t,v}) \right]. \quad (19)$$

On peut appliquer le schéma de Monte Carlo sur les seules variables (t, v) de la relation (19).

6.3.2 Exercice

Pour résoudre le problème numériquement en s'appuyant directement sur (18), on utilise le schéma de la Définition 6.4. En introduisant $\Delta t > 0$ et $\mathbf{T}^{\Delta t}$, et le contrôle $\phi^{a^*} = a^* \in \mathbb{R}$ constant, il s'agit de résoudre

$$\begin{cases} v_\circ(t, v) = \sup_{a^* \in \mathbb{R}} \mathbb{E} \left[e^{-\eta X_{t+\Delta t}^{t,0,v,\phi^{a^*}}} v_\circ(t + \Delta t, V_{t+\Delta t}^{t,v}) \right] & (t, v) \in \mathbf{T}^{\Delta t} \setminus \{T\} \times \mathbf{V} \\ v_\circ(T, v) = U(0) = 1 & v \in \mathbf{V} \end{cases} \quad (20)$$

où il faudra à nouveau appliquer le barycentre sur $V_{t+\Delta t}^{t,v}$, ce dernier ne sera pas en général dans \mathbf{V} .

- Résolvez numériquement (20) le problème discret associé, c'est-à-dire, calculez pour tout $t \in [0, T]$ et pour tout $\sqrt{v} \in \sqrt{\mathbf{V}}$ (défini ci-après), la fonction $v_\circ(t, v)$ et le contrôle optimal $\hat{\phi}$ (qui est une fonction de $[t, v]$). On utilisera
 - $\eta = 0.1$,
 - $\rho = -0.7$,
 - $\mu = 0.07$,
 - $\sigma^2 = 0.04$,
 - $\alpha = 2$,
 - $\xi = 0.25$,
 - $T = 1$,
 - $\Delta t = 0.004$ (251 points),
 - $\sqrt{\mathbf{V}} = \{0.01, 0.01 + \Delta v, 0.01 + 2\Delta v, \dots, 0.51\}$ avec $\Delta v = 0.005$ (101 points).

On testera une allocation $a^* \in [0, 100]$ puis on calculera, avec au moins 1000 simulations, $J(t, v, a^*)$, et on en déduira $\hat{\phi}[t, v]$ optimal et $v_\circ(t, v) = J(t, v, \hat{\phi}[t, v])$. On se restreindra à $a^* \in \{0, 1, 2, \dots, 100\}$, on pourra tous les tester et en déduire celui qui atteint le maximum.

Remarque 1. On rappelle que les simulations de $V_{t+\Delta t}^{t,v}$ ne tomberont pas dans \mathbf{V} . Dans ce cas là, on utilise le barycentre pour calculer la valeur (interpolation linéaire). Ceci est fondamental et prendre le point le plus proche biaiserait le résultat : si le pas de temps est très petit, il se peut que $V_{t+\Delta t}^{t,v}$ soit toujours très proche de v ce qui ferait considérer au problème V comme constant (ou quasi constant).

En cas de sortie du maillage, on prendra la valeur la plus proche.

Remarque 2. Pour améliorer l'estimation de $J(t, v, a^*)$, on utilisera les variables aléatoire anthétiques.

- Illustrez en

- Affichant la fonction de contrôle optimal $\hat{\phi}[t, v]$;
- Simulant (Y_t^{0, Y_0, V_0}) et (V_t^{0, V_0}) et en affichant les trajectoires de (Y_t^{0, Y_0, V_0}) , de (V_t^{0, V_0}) , de $\hat{\phi}[t, V_t^{0, V_0}]$ et de (X_t^{0, X_0, V_0}) en partant d'une richesse initiale de $X_0 = 100$, d'un prix initial de $Y_0 = 10$ et d'une variance initiale de $V_0 = 0.04$.

6.4 Schéma explicite

Nous reprenons $\partial_{\Delta t}\varphi(t, x)$, $\partial_{\Delta x}^+\varphi(t, x)$ et $\partial_{\Delta x}^-\varphi(t, x)$ définies dans (11).

De l'équation (15), on déduit le schéma explicite qui consiste à résoudre de manière rétrograde en t

$$\partial_{\Delta t}\varphi(t, x) + \sup_{a \in A} \left[\mu(t, x, a) \partial_{\Delta x}\varphi(t + \Delta t, x) + \frac{1}{2} \sigma^2(t, x, a) \partial_{\Delta x}^2\varphi(t + \Delta t, x) \right] = 0$$

avec

$$\mu(t, x, a) \partial_{\Delta x}\varphi(t + \Delta t, x) := \begin{cases} \mu(t, x, a) \partial_{\Delta x}^+\varphi(t + \Delta t, x) & \text{si } \mu(t, x, a) \geq 0, \\ \mu(t, x, a) \partial_{\Delta x}^-\varphi(t + \Delta t, x) & \text{si } \mu(t, x, a) < 0. \end{cases}$$

et partant de $\varphi(T, x) = g(x)$. On se restreint à un ensemble compact $\mathbf{A} \subset A$.

Définition 6.5 (Schéma par Différences finies).

- On pose $\varphi(T, x) = g(x)$ pour $x \in \mathbf{X}^{\Delta x}$,
- Sachant $\varphi(t + \Delta t, \cdot)$ pour un $t \in \mathbf{T}^{\Delta t}$, on déduit de l'équation des différences finies :

$$\hat{a}[t, x] \in \arg \max_{a \in \mathbf{A}} \left[\mu(t, x, a) \partial_{\Delta x}\varphi(t + \Delta t, x) + \frac{1}{2} \sigma^2(t, x, a) \partial_{\Delta x}^2\varphi(t + \Delta t, x) \right]$$

et

$$\begin{aligned} \varphi(t, x) &= \left(1 - \sigma^2(t, x, \hat{a}[t, x]) \frac{\Delta t}{\Delta x^2} - |\mu(t, x, \hat{a}[t, x])| \frac{\Delta t}{\Delta x} \right) \varphi(t + \Delta t, x) \\ &+ \left(\frac{1}{2} \sigma^2(t, x) \frac{\Delta t}{\Delta x^2} + \mu^+(t, x, \hat{a}[t, x]) \frac{\Delta t}{\Delta x} \right) \varphi(t + \Delta t, x + \Delta x) \\ &+ \left(\frac{1}{2} \sigma^2(t, x) \frac{\Delta t}{\Delta x^2} + \mu^-(t, x, \hat{a}[t, x]) \frac{\Delta t}{\Delta x} \right) \varphi(t + \Delta t, x - \Delta x). \end{aligned}$$

Comme dans le cas sans contrôle, le schéma devra vérifier la condition de monotonie.

Remarque 6.6. Si $(\Delta t, \Delta x) \rightarrow (0, 0)$ et tel qu'on ait toujours, pour $(\Delta t, \Delta x)$ suffisamment petit

$$1 - \sigma^2(t, x, \hat{a}[t, x]) \frac{\Delta t}{\Delta x^2} - |\mu(t, x, \hat{a}[t, x])| \frac{\Delta t}{\Delta x} \geq 0,$$

la condition de monotonie sera satisfaite.

Si on arrive à montrer que la famille de solution reste bornée sur tout compact, alors le schéma est convergent lorsque $(\Delta t, \Delta x) \rightarrow (0, 0)$ si la condition de monotonie est satisfaite.

6.5 Exemple - Gestion de portefeuille (suite)

Nous reprenons l'exemple de la section 6.3 que nous résolvons, cette fois-ci, avec un schéma numérique sur l'EDP associée.

On rappelle l'équation (17) : la fonction valeur v vérifie

$$v(t, x, v) = e^{-\eta x} v(t, 0, v),$$

et que nous avons posé $v_\circ(t, v) := v(t, 0, v)$. Nous pourrions partir de (19), et retrouver (en appliquant Itô) l'équation aux dérivées d'Hamilton Jacobi Bellman vérifiée par v_\circ . Mais nous pouvons également remarquer de (17) que

$$\partial_x v(t, x, v) = -\eta v(t, x, v) \quad (21)$$

$$\partial_{xx}^2 v(t, x, v) = \eta^2 v(t, x, v) \quad (22)$$

$$\partial_{xv}^2 v(t, x, v) = -\eta \partial_v v(t, x, v) \quad (23)$$

De (6.1) on déduit l'équation aux dérivées partielle vérifiée par v (version en dimension 2)

$$\partial_t v(\cdot) + \sup_{a \in A} \left[a\mu \partial_x v(\cdot) + \frac{1}{2} a^2 v \partial_{xx}^2 v(\cdot) - \alpha(v - \sigma^2) \partial_v v(\cdot) + \frac{1}{2} \xi^2 v \partial_{vv}^2 v(\cdot) + a\rho\xi v \partial_{xv}^2 v(\cdot) \right] = 0$$

Dans l'équation ci-dessus, nous injectons les relations (21) pour les dérivées partielles par rapport à x .

$$\partial_t v(\cdot) - \alpha(v - \sigma^2) \partial_v v(\cdot) + \frac{1}{2} \xi^2 v \partial_{vv}^2 v(\cdot) + \sup_{a \in A} \left[-a\mu\eta v(\cdot) + \frac{1}{2} a^2 v \eta^2 v(\cdot) - a\rho\xi v \eta \partial_v v(\cdot) \right] = 0$$

En multipliant par $e^{\eta x}$ l'équation ci-dessus, on en déduit que la même est vérifiée pour v_\circ c'est à dire

$$\partial_t v_\circ(\cdot) - \alpha(v - \sigma^2) \partial_v v_\circ(\cdot) + \frac{1}{2} \xi^2 v \partial_{vv}^2 v_\circ(\cdot) + \sup_{a \in A} \left[-a\mu\eta v_\circ(\cdot) + \frac{1}{2} a^2 v \eta^2 v_\circ(\cdot) - a\rho\xi v \eta \partial_v v_\circ(\cdot) \right] = 0$$

Par définition, v_\circ est une fonction strictement négative. Le sup ci-dessus est atteint dans un unique point sur \mathbb{R} . Soit $\delta < 0$ et $\beta \in \mathbb{R}$. On a $\arg \max_{a \in \mathbb{R}} \{ \frac{1}{2} \delta a^2 - \beta a \}$ qui est atteint en $a = \frac{\beta}{\delta}$ et le max vaut $\frac{-\beta^2}{2\delta}$. On en déduit le contrôle optimal :

$$\hat{a}[t, v] = \frac{\mu\eta v_\circ(t, v) + \rho\xi v \eta \partial_v v_\circ(t, v)}{v\eta^2 v_\circ(t, v)} = \frac{\mu v_\circ(t, v) + \rho\xi v \partial_v v_\circ(t, v)}{v\eta v_\circ(t, v)}$$

On pose

$$H(t, v) := -\frac{(\mu v_\circ(t, v) + \rho\xi v \partial_v v_\circ(t, v))^2}{2v v_\circ(t, v)}$$

L'EDP sur v_\circ devient :

$$\partial_t v_\circ(t, v) - \alpha(v - \sigma^2) \partial_v v_\circ(t, v) + \frac{1}{2} \xi^2 v \partial_{vv}^2 v_\circ(t, v) + H(t, v) = 0$$

Avec $H_{\Delta v}$ défini de manière analogue à H pour le schéma explicite, celui-ci devient :

$$\partial_{\Delta t} \varphi(t, v) - \alpha(v - \sigma^2) \partial_{\Delta v} \varphi(t + \Delta t, v) + \frac{1}{2} \xi^2 v \partial_{\Delta v}^2 \varphi(t + \Delta t, v) + H_{\Delta v}(t + \Delta t, v) = 0$$

et le contrôle optimal, qui peut être déduit après résolution de l'équation de la fonction valeur, peut être estimé par

$$\hat{a}[t, v] = \frac{\mu\varphi(t, v) + \rho\xi v \partial_{\Delta v} \varphi(t, v)}{v\eta\varphi(t, v)}$$

6.5.1 Exercice

Il s'agit de résoudre en utilisant le schéma explicite, c'est à dire appliquer l'algorithme

$$\begin{cases} \varphi(t, v) = \varphi(t + \Delta t, v) - \alpha(v - \sigma^2) \partial_{\Delta v} \varphi(t + \Delta t, v) \Delta t \\ \quad + \frac{1}{2} \xi^2 v \partial_{\Delta v}^2 \varphi(t + \Delta t, v) \Delta t + H_{\Delta v}(t + \Delta t, v) \Delta t & (t, v) \in \mathbf{T}^{\Delta t} \setminus \{T\} \times \mathbf{V} \\ \varphi(T, v) = U(0) = 1 & v \in \mathbf{V} \end{cases} \quad (24)$$

Résolvez numériquement avec les mêmes paramètres que l'exercice en 6.3.2.

A Simulations - Méthode du rejet

Lemme A.1. Soit g une densité de probabilité et $c > 0$ une constante. Alors l'algorithme :

- Tirer X suivant la densité g ,
- Tirer U suivant la loi uniforme sur $[0, cg(X)]$,

permet de tirer (X, U) suivant la loi uniforme sur l'ensemble

$$\mathcal{A} = \{(x, u) \in \mathbb{R}^2, \text{ t.q. } 0 \leq u \leq cg(x)\}.$$

Démonstration. En notant λ_2 la mesure de Lebesgue sur \mathbb{R}^2 , la loi uniforme sur \mathcal{A} a la densité :

$$\frac{1}{\lambda_2(\mathcal{A})} \mathbf{1}_{\mathcal{A}}.$$

On remarque que :

$$\lambda_2(\mathcal{A}) = \int_{\mathbb{R}} \int_{\mathbb{R}} \mathbf{1}_{\{0 \leq u \leq cg(x)\}}(x, u) du dx = \int_{\mathbb{R}} cg(x) dx = c.$$

D'autre part, en notant $f_{(X,U)}$ la densité sur \mathbb{R}^2 du couple (X, U) et $f_{U|\{X=x\}}$ la densité conditionnelle de $U | \{X = x\}$, on a (là où $g(x) > 0$, sinon le résultat est trivial) :

$$f_{(X,U)}(x, u) = f_X(x) f_{U|\{X=x\}}(u) = g(x) \frac{1}{cg(x)} \mathbf{1}_{\{0 \leq u \leq cg(x)\}} = c^{-1} \mathbf{1}_{\mathcal{A}}.$$

□

Lemme A.2. Soient f et g deux densités de probabilité et $c \geq 1$ une constante telle que $f \leq cg$. Soit $(X, U) \sim \mathcal{U}(\mathcal{A})$. On introduit

$$\mathcal{B} = \{(x, u) \in \mathbb{R}^2, \text{ t.q. } 0 \leq u \leq f(x)\}.$$

Alors

$$(X, U) | U \leq f(X) \sim \mathcal{U}(\mathcal{B}).$$

Démonstration. Dans un premier temps :

$$\begin{aligned} \mathbb{P}(U \leq f(X)) &= c^{-1} \int_{\mathbb{R}} \int_{\mathbb{R}} \mathbf{1}_{\{u \leq f(x)\}} \mathbf{1}_{\{0 \leq u \leq cg(x)\}} du dx \\ &= c^{-1} \int_{\mathbb{R}} \int_{\mathbb{R}} \mathbf{1}_{\{0 \leq u \leq f(x)\}} du dx = c^{-1} \int_{\mathbb{R}} f(x) dx = c^{-1}. \end{aligned}$$

On en déduit, pour tout borélien $A \in \mathbb{B}(\mathbb{R}^2)$,

$$\begin{aligned} \mathbb{P}[(X, U) \in A | U \leq f(X)] &= \frac{\mathbb{P}[(X, U) \in A, U \leq f(X)]}{\mathbb{P}[U \leq f(X)]} \\ &= \int_A \mathbf{1}_{\{u \leq f(x)\}} \mathbf{1}_{\{0 \leq u \leq cg(x)\}} du dx = \int_A \mathbf{1}_{\{0 \leq u \leq f(x)\}} du dx. \end{aligned}$$

□

Proposition A.3 (Méthode du rejet). Soient f et g deux densités de probabilité et $c \geq 1$ une constante telle que $f \leq cg$. La variable aléatoire issue de l'algorithme du rejet :

- (1) Tirer X suivant la densité g ,
- (2) Tirer U suivant la loi uniforme sur $[0, cg(X)]$,
- (3) Si $U \leq f(X)$, conserver X , sinon recommencer à l'étape (1).

a pour densité la fonction f .

Démonstration. On remarque que par le lemme A.1, à l'issue des étapes (1) et (2), on a par construction $(X, U) \sim \mathcal{U}(\mathcal{A})$. L'étape (3) n'accepte X que si le couple (X, U) est dans \mathcal{B} , sinon on recommence, ce qui revient à simuler uniformément dans \mathcal{B} par le lemme A.2. En vertu du lemme A.1, X a alors la loi de probabilité de densité f (cas particulier $c = 1$). \square

Proposition A.4. *Le nombre N nécessaire de simulations de couples (X, U) afin de conserver une simulation suit une loi géométrique qui démarre à 1 et de paramètre c^{-1} , c'est-à-dire :*

$$\mathbb{P}(N = n) = c^{-1}(1 - c^{-1})^{n-1}, \quad n \geq 1.$$

En particulier,

$$\mathbb{E}(N) = c.$$

Démonstration. Soit $(X_n, U_n)_{n \geq 1}$ une suite i.i.d. de v.a.r. de loi $\mathcal{U}(\mathcal{A})$, par construction :

$$N = \min\{n \geq 1 \mid (U_n, X_n) \in \mathcal{B}\}.$$

On a par construction :

$$\{N = n\} = \{(U_1, X_1) \notin \mathcal{B}, \dots, (U_{n-1}, X_{n-1}) \notin \mathcal{B}, (U_n, X_n) \in \mathcal{B}\}.$$

Par indépendance,

$$\begin{aligned} \mathbb{P}(N = n) &= \mathbb{P}((U_1, X_1) \notin \mathcal{B}, \dots, (U_{n-1}, X_{n-1}) \notin \mathcal{B}, (U_n, X_n) \in \mathcal{B}) \\ &= \left(\prod_{i=1}^{n-1} \mathbb{P}((U_i, X_i) \notin \mathcal{B}) \right) \mathbb{P}((U_n, X_n) \in \mathcal{B}) \\ &= c^{-1}(1 - c^{-1})^{n-1}. \end{aligned}$$

On a utilisé ci-dessus le fait que, pour tout $i \geq 1$, $\mathbb{P}((U_i, X_i) \in \mathcal{B}) = \mathbb{P}(U_i \leq f(X_i)) = c^{-1}$ qui a été établi dans la démonstration du lemme A.2. \square

Exercice 16.

— Montrez que, pour $x \geq 0$,

$$\frac{\sqrt{2}}{\sqrt{\pi}} e^{-\frac{1}{2}x^2} \leq ce^{-x},$$

avec $c \geq \frac{e^{\frac{1}{2}}\sqrt{2}}{\sqrt{\pi}}$,

— En déduire une méthode du rejet pour la simulation de $|X|$ où $X \sim \mathcal{N}(0, 1)$ et l'implémenter sous Python.

— Montrez que le test de rejet peut s'écrire $\log(U) \leq -\frac{1}{2}(X - 1)^2$ avec $U \sim \mathcal{U}([0, 1])$, implémentez ce test et comparez le temps de calcul.

On utilisera `random.exponential` du package `numpy` pour simuler des lois exponentielles (`rexp` en R). Pour comparer le temps de calcul on utilisera la fonction `time` du package `time` (`proc.time` en R) qui donne le temps écoulé en secondes depuis 1970.

B Formule de Black Scholes

$$\begin{aligned}d_1 &:= \frac{1}{\sigma\sqrt{T}} [\log(S_0/K) + (r + \sigma^2/2) T] \\d_2 &:= d_1 - \sigma\sqrt{T} \\C &= S_0\Phi(d_1) - Ke^{-rT}\Phi(d_2) \\P &= -S_0\Phi(-d_1) + Ke^{-rT}\Phi(-d_2)\end{aligned}$$

C Exercice 11 - Prix des Zéros Coupon dans le modèle de Vašíček

La valeur exacte de $P(t, T)$ dans le modèle peut se calculer et est donnée par

$$P(t, T) = e^{A(t, T) - B(t, T)r_t},$$

avec

$$\begin{aligned}A(t, T) &:= \left(\mu - \frac{\sigma^2}{2a^2}\right) (B(t, T) - (T - t)) - \frac{\sigma^2}{4a} B(t, T)^2, \\B(t, T) &:= \frac{1 - e^{-a(T-t)}}{a}.\end{aligned}$$

D Exercice 14 - Option asiatique - Calcul de l'espérance de X

Dans cet exercice, pour appliquer la méthode de réduction de variance avec X , il faut calculer $\mathbb{E}(X)$ avec

$$X := \left(\exp\left(\frac{1}{T} \int_0^T \log(S_t) dt\right) - K \right)^+.$$

Comme

$$\log(S_t) = \log(S_0) + (r - \sigma^2/2)t + \sigma W_t,$$

il vient

$$\frac{1}{T} \int_0^T \log(S_t) dt = \log(S_0) + (r - \sigma^2/2) T/2 + \frac{\sigma}{T} \int_0^T W_t dt.$$

On peut montrer que

$$\int_0^T W_t dt \sim \mathcal{N}\left(0, \frac{T^3}{3}\right).$$

Pour cela on approxime l'intégrale par des sommes de Riemann et on observe qu'il s'agit d'une limite L^2 de variable aléatoires gaussiennes centrées. Puis on calcule la limite de l'espérance du carré avec cette même approximation.

Avec $\bar{\sigma}^2 := \sigma^2/3$ et $W'_T \sim \mathcal{N}(0, T)$, on a en loi

$$\begin{aligned}
\exp\left(\frac{1}{T}\int_0^T \log(S_t)dt\right) &= S_0 \exp\left((r - \sigma^2/2)T/2 + \bar{\sigma}W'_T\right) \\
&= S_0 \exp\left((r/2 - \sigma^2/4) + \bar{\sigma}^2/2 - \bar{\sigma}^2/2\right)T + \bar{\sigma}W'_T \\
&= S_0 \exp\left((r/2 - \sigma^2/12) - \bar{\sigma}^2/2\right)T + \bar{\sigma}W'_T.
\end{aligned}$$

L'espérance de X , au facteur d'actualisation $e^{\bar{r}T}$ près, s'obtient avec la formule de Black Sholes du Call, en prenant les nouveaux paramètres $(\bar{r}, \bar{\sigma})$ avec

$$\bar{r} := r/2 - \sigma^2/12.$$

C'est-à-dire

$$\mathbb{E}(X) = e^{\bar{r}T} (S_0\Phi(d_1) - Ke^{-\bar{r}T}\Phi(d_2)).$$

où d_1 et de d_2 sont définis dans l'annexe B, et sont évalués avec \bar{r} et $\bar{\sigma}$.